《俄罗斯数学教材选译》序

从上世纪 50 年代初起, 在当时全面学习苏联的大背景下, 国内的高等学校大量采用了翻译过来的苏联数学教材. 这些教材体系严密, 论证严谨, 有效地帮助了青年学子打好扎实的数学基础, 培养了一大批优秀的数学人才. 到了 60 年代, 国内开始编纂出版的大学数学教材逐步代替了原先采用的苏联教材, 但还在很大程度上保留着苏联教材的影响, 同时, 一些苏联教材仍被广大教师和学生作为主要参考书或课外读物继续发挥着作用. 客观地说, 从解放初一直到文化大革命前夕, 苏联数学教材在培养我国高级专门人才中发挥了重要的作用, 起了不可忽略的影响, 是功不可没的.

改革开放以来, 通过接触并引进在体系及风格上各有特色的欧美数学教材, 大家眼界为之一新, 并得到了很大的启发和教益. 但在很长一段时间中, 尽管苏联的数学教学也在进行积极的探索与改革, 引进却基本中断, 更没有及时地进行跟踪, 能看懂俄文数学教材原著的人也越来越少, 事实上已造成了很大的隔膜, 不能不说是一个很大的缺憾.

事情终于出现了一个转折的契机. 今年初, 在由中国数学会、中国工业与应用数学学会及国家自然科学基金委员会数学天元基金联合组织的迎春茶话会上, 有数学家提出, 莫斯科大学为庆祝成立 250 周年计划推出一批优秀教材, 建议将其中的一些数学教材组织翻译出版. 这一建议在会上得到广泛支持, 并得到高等教育出版社的高度重视. 会后高等教育出版社和数学天元基金一起邀请熟悉俄罗斯数学教材情况的专家座谈讨论, 大家一致认为: 在当前着力引进俄罗斯的数学教材, 有助于扩大视野, 开拓思路, 对提高数学教学质量、促进数学教材改革均十分必要. 《俄罗斯数学教材选译》系列正是在这样的情况下, 经数学天元基金资助, 由高等教育出版社组织出版的.

经过认真选题并精心翻译校订, 本系列中所列入的教材, 以莫斯科大学的教材为主, 也包括俄罗斯其他一些著名大学的教材. 有大学基础课程的教材, 也有适合大学高年级学生及研究生使用的教学用书. 有些教材虽曾翻译出版, 但经多次修订重版, 面目已有较大变化, 至今仍广泛采用、深受欢迎, 反射出俄罗斯在出版经典教材方面所作的不懈努力, 对我们也是一个有益的借鉴. 这一教材系列的出版, 将中俄数学教学之间中断多年的链条重新连接起来, 对推动我国数学课程设置和教学内容的改革, 对提高数学素养、培养更多优秀的数学人才, 可望发挥积极的作用, 并起着深远的影响, 无疑值得庆贺, 特为之序.

李大潜

2005 年 10 月

序

专门出版这本“概率论习题集”的想法是这样产生的.

在我们的书《概率》的前两版 (1980, 1989) 中, 全书八章的每一章中都配备有大量的各种类型的习题. 在准备出版第三版时, 由于改写和增补的部分较多, 将书分为了两卷, 于 2004 年 4 月份出版. 也正是在此过程中产生了单独出版一本习题集, 以把原书中的“老习题”, 以及由于各种原因未能收入原书中的“新习题”都包括进去的想法. (未能收入原书的一个基本原因是受到在书籍出版时关于篇幅限制的约束.)

习题的收集和剔除是我在多年之中根据自己的兴趣进行的, 习题的来源多种多样, 各种教科书、讲义、习题集、专著、刊物上面发表的论文, ……. 还有一些题目产生于我们为本科生和研究生举办的专门的讨论班.

很难逐一列出所有习题的确切来源. 但是其中不少可以从本书最后所附的参考文献中反映出来.

绝大多数习题后面都给出了提示, 这样做的原因如下:

我们的《概率》书诞生在国立莫斯科大学数学力学系, 是在我们的讲座的基础上形成的, 那里的本科生在第四学期末选择专业. 那些在三年级时以我为导师的学生会被告知: 你们 (在三年级) 的第一项工作由两部分组成, 其中的第一部分就是解答《概率》书中的习题, 例如第二卷第八至十章中的习题. 那么毫无疑问, 学生们为了解答这些习题, 不仅需要独立熟悉相应章节的内容, 而且在必要时, 还要熟悉以前学过的有关章节. 在收到他们的解答 (不是手写的, 而是用 Tex 打出来的) 之后, 通过批阅和讨论这些解答, 学生们获得了他们的第二部分工作, 这里面就已经有着更多的创造性了.

这样多年坚持的结果, 使我们积累起数量可观的对各种题目 (收录在《概率》的

各个版本中的题目) 的解答, 其中的多数解答曾按照我的要求反复进行过验证、校勘, 并且幸运地被我们教研室中的一些合作者、研究生和助教们系统地重新做解, 他们之中有 (按照解答的章节顺序排列): А. С. Черного (第一至五章), М. А. Урусов, И. А. Ярошенков 和 Ю. А. Кузнецов (第六至八章). 我谨在此对他们的工作致以深切的谢意.

收集完所有习题解答之后, 我立即意识到不可能立即出版它们, 因为它们所占的篇幅实在太大. 我不得不将习题解答改写为提示, 正如现在的这种形式, 此前我还为数目众多的新题目写了提示.

应当特别说一说放在本书末尾的附录. 通过附录, 我们可以较早地把一些概念介绍给读者, 尽管只能是初浅的介绍, 这样做的理由如下. 一方面, 附录中对本书以及《概率》第一、二卷中的一些基本符号、基础性概念给出了简介性、评述性的提示; 另一方面, 附录中收入了一些解题时需要用到的组合论、马尔可夫链的位势理论方面的补充材料, 其中有些概念是我们书的正文中所没有出现过的.

现在的这本习题集是与我们的《概率》第一、二卷紧密联系的. 例如, 在本书第二章第 3 节第 11 题的文字中写道: “试验证表 2 和表 3 中所列入的‘分布’都是真正的概率分布”. 应当明白, 这里所说的表 2 和表 3 是指《概率》第一卷第二章第 3 节中的表 2 和表 3 (p162–163). 又如, 第四章第 3 节中的题目 5(c):“证明, (恰姆别尔诺伊命题中的) 数 $\omega = 0.123456789101112\cdots$, 其中依次所写出的所有的数都是‘正常的’(按十进制表示, 参阅例 2)”. 这里所说的例 2 是在《概率》第二卷第四章第 3 节中 (p19).

读者们将会发现, 收集在本书中的习题有着多种不同的性质.

(A) 一些题目和练习是用来检查对《概率》第一、二卷中所介绍过的概念、事实和结论的掌握情况的 (例如: 第一章第 1, 2 节关于有利场合下的样本点个数的组合计数, 以及为此目的而展开的各种概念, 诸如非完全阶乘 $(N)_n$, 组合数 C_N^n 和 C_{N+n-1}^n, 卡塔兰数 C_n, 一阶和二阶斯特林数 s_N^n 和 S_N^n, 贝尔数 B_n, 斐波拉契数 F_n, 等等).

(B) 另外一些 (中等难度和高等难度的) 习题则要求具有较大的创造性 (例如: 证明将关于强收敛的勒贝格定理和关于条件数学期望中的极限性状的莱维定理联系起来的结果, 见第七章第 4 节第 3 题).

(C) 为了陈述许多题目, 需要补充《概率》第一、二卷正文中所没有出现的一些资料, 介绍读者认识一些已经熟知的或者关于其存在性已经有所证实的事实 (例如, 第二章第 2 节第 27 题介绍了苏斯林的如下结论: 平面博雷尔集在一个坐标轴上的投影可以不是直线上的博雷尔集; 还介绍了对集合进行怎样的运算, 可以得到由某个集合类所生成的最小代数和 σ-代数中的所有集合, 参阅第二章第 2 节第 25, 26 和 32 题).

(我们完全清楚, 这种类型的题目事实上就是具有难度的定理. 然而, 把它们写成习题的形式可以引导读者去思索, 例如, 在非初等的概率论中, 考虑如何可以实际

地构造出在概率模型的描述中起着关键作用的完整的 σ-代数来.)

(D) 许多题目与随机游动朝向布朗运动和布朗桥收敛的极限过程有关 (例如第三章第 4 节的习题). 这些关于 "不变原理" 的习题是认识连续时间随机过程, 特别地, 是认识泛函极限定理中的问题的前奏.

再谈一些其他方面的问题.

多年以来, 莫斯科大学的教师们曾出版过若干本概率论习题集, 它们不仅常年被使用在莫斯科大学的教学过程中, 而且还被许多其他高等院校采用. 它们是:

1963 年, Мешалкин Л. Д. 概率论习题集, 莫斯科大学出版社;

1980 年, Севастьянов Б. А., Чистяков В. П., Зубков А. М. 概率论习题集, 莫斯科: 科学出版社;

1986 年, Прохоров А. В., Ушаков В. Г., Ушаков Н. Г. 概率论习题集 (基本概念. 极限定理. 随机过程), 莫斯科: 科学出版社;

1989 年, Зубков А. М., Севастьянов Б. А., Чистяков В. П. 概率论习题集, 第二版, 莫斯科: 科学出版社;

1990 年, Козлов М. В. 初等概率论中的例题与习题, 莫斯科大学出版社.

自从上述习题集中的最后一本出版以来, 已经十五年过去了. 在这段时间中, 概率论教科书的结构已经发生了很大的变化. 出现了许多新的方向、新的分支和新的题目. 非常希望我们这本于 2006 年出版的习题集, 能够与上列诸本习题集一道, 不仅能够充分反映出概率论的现状, 而且能够充分反映出它的传统和经典的部分.

在结束序言之时, 我们指出, 本书共约收录了一般难度的习题 1500 道 (包括子题在内).

正如在出版《概率》第一、二卷时那样, 在本书的出版准备工作中, 托洛佐娃 (Т. Б. Толозова) 一如既往地做了大量的校对、录入文字和学术编辑工作, 谨在此向她致以诚挚的谢意.

A. 施利亚耶夫
莫斯科大学概率论教研室
俄罗斯科学院数学研究所
莫斯科, 2004 年

目 录

第一章 初等概率论

§1. 有限种结局试验的概率模型

1. 考察对于集合 Ω 的子集进行的博雷尔交 $(\cap)$ 并 $(\cup)$ 运算, 验证它们具有下述性质:

$A\cup B=B\cup A,\quad A\cap B=B\cap A$ (交换律),

$A\cup(B\cup C)=(A\cup B)\cup C,\quad A\cap(B\cap C)=(A\cap B)\cap C$ (结合律),

$A\cap(B\cup C)=(A\cap B)\cup(A\cap C),\ A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$ (分配律),

$A\cup A=A,\quad A\cap A=A$ (幂等性).

并且证明, 对于交并运算的取余运算, 还有如下的德摩根法则成立:

$$\overline{A\cup B}=\overline{A}\cap\overline{B},\quad \overline{A\cap B}=\overline{A}\cup\overline{B}$$

(其中 $\overline{A}$ 表示集合 A 的补集, 即 $\Omega\backslash A$).

2. (不完全阶乘 $(N)_n$ 的性质与各种解释, 其中 $(N)_n\equiv N(N-1)\cdots(N-n+1)$ 是自 N 个元素中取出 n 个来的排列数, 具体介绍参阅附录第 1 节.) 证明,

(a) 由共有 $|A|=N$ 个元素的集合 A 中取出的大小为 n 的有序样本 $(\cdots)$(无重复元素, 即 "无放回抽样") 的数目等于 $(N)_n$, $1\leqslant n\leqslant N$.

(b) 从由 N 个不同字母构成的字母表中取出不同字母所形成的长度为 n 的单词的数目为 $(N)_n$, $1\leqslant n\leqslant N$.

(c) 定义在集合 X $(|X|=n)$ 上, 取值于集合 Y $(|Y|=N, N\geqslant n)$, 并且满足性质 "如果 $x_1\neq x_2$, 则 $f(x_1)\neq f(x_2)$" 的不同函数 $y=f(x)$(即 $f:X\to Y$ 单射) 的数目等于 $(N)_n$.

3. (二项系数 $C_N^n \equiv \frac{N!}{n!(N-n)!}$ 的性质与各种解释, 具体介绍参阅附录第 1 节.) 证明,

(a) 由共有 $|A|=N$ 个元素的集合 A 中取出的大小为 n 的无序样本 $[\cdots]$(无重复元素, 即 "无放回抽样") 的数目等于 C_N^n, $1 \leqslant n \leqslant N$.

(b) 由 n 个 1 和 $N-n$ 个 0 所组成的有序数组 $(\cdots)$ 的数目等于 C_N^n, $1 \leqslant n \leqslant N$.

(c) n 个不可区分的小球放入 N 个不同的盒子, 每个盒子至多放 1 个小球 ("有限制的放球") 的放法数目等于 C_N^n, $1 \leqslant n \leqslant N$.

(d) 2 维非负格点集合 $\mathbb{Z}_+^2=\{(i,j):\ i,j=0,1,2,\cdots\}$ 上的以原点 $(0,0)$ 为起点, 以 $(n,N-n)$ 为终点的非降路径(所谓非降路径, 就是每一步都只能向上或者向右移动一个单位的路径) 的条数等于 C_N^n, $0 \leqslant n \leqslant N$, $C_N^0=1$.

(e) 元素个数 $|A|=N$ 的集合 A 的由 n 个元素构成的不同的子集 D ($|D|=n \leqslant N$) 的数目等于 C_N^n.

提示: 如果首先直接证明其中的一个命题, 例如先证命题 (a) 成立, 那么其余命题就可以利用已经证得的结果来证明, 例如可采用证明途径: (a)$\rightleftharpoons$(b), (a)$\rightleftharpoons$(c), $\cdots$, 参阅第 1 节例 6.

4. 如同上题中的 (d), 我们考察 2 维非负格点集合 $\mathbb{Z}_+^2=\{(i,j):\ i,j=0,1,2,\cdots\}$ 上的以原点 $(0,0)$ 为起点, 以 (n,n) 为终点的位于 "对角线" 下方, 且至多与之相切的路径 (即仅经过集合 $\{(i,j)\in\mathbb{Z}_+^2:\ 1\leqslant j\leqslant i\leqslant n\}$ 中的点的路径).

证明, 这样的路径的条数等于 C_{n+1}, 其中

$$C_n=\frac{1}{n}C_{2(n-1)}^{n-1}, \quad n \geqslant 1,$$

称为卡塔兰 (Catalan) 数, 见 [45]. (有时也将卡塔兰数理解为 $c_n=C_{n+1}$, $n\geqslant 1$, 例如, 参阅 [61].)

试验证 $C_1,\cdots,C_9$ 的值分别为 $1,1,2,5,14,42,132,429,1430$.

5. 对卡塔兰数 C_n, $n\geqslant 1$, 有许多有趣的组合解释. 例如, 对于 n 个数 $a,b,c,d,\cdots$ 的求和, 如果每次只将两个数相加 (但不允许改变加数的顺序), 我们来考察不同的求和方式的数目 N_n. 比如说, 当 $n=3$ 时, 如果按照每次只将两个数相加的规则, 和 $a+b+c$ 可以有如下各种不同的 "加括号" 方式: $a+b+c=\Big(a+(b+c)\Big)=\Big((a+b)+c\Big)$, 此处不同的求和方式数目 $N_3=2$. 对于 $n=4$, 我们有 5 种不同的 "加括号" 方式 ($N_4=5$):

$$\begin{aligned} a+b+c+d &= \Big((a+b)+(c+d)\Big)=\Big(((a+b)+c)+d\Big) \\ &= \Big((a+(b+c))+d\Big)=\Big(a+((b+c)+d)\Big)=\Big(a+(b+(c+d))\Big). \end{aligned}$$

(a) 证明, 对任何 $n\geqslant 3$, 不同的求和方式数目 N_n 都等于 C_n.

(b) 我们来观察正 n 边形 ($n\geqslant 4$) 的用其对角线所作的三角形剖分, 即用不在形内相交的对角线将其分为一系列三角形 (显然, 自每个顶点所引出的对角线条数等于 $n-3$). 证明, 这样的剖分数目 N_n 等于 C_{n-1}.

(c) 证明, 卡塔兰数 C_n, $n>1$, 满足递推关系式

$$C_n^*=\sum_{i=1}^{n-1}C_i^*C_{n-i}^*, \qquad (*)$$

其中 $C_0^*=0$, $C_1^*=1$.

(d) 证明, 由递推式 $(*)$ 所定义的序列 $(C_n^*)_{n\geqslant 1}$ 的母函数 $F^*(x)=\sum\limits_{n\geqslant 1}C_n^*x^n$ 满足方程

$$F^*(x)=x+\Big(F^*(x)\Big)^2.$$

(e) 证明 (注意 $F^*(0)=0$),

$$F^*(x)=\frac{1}{2}\Big(1-(1-4x)^{1/2}\Big), \quad |x|<\frac{1}{4}.$$

并由此算出 (x^n 的) 系数 C_n^*, 正如我们所期望的, 它重合于卡塔兰数 C_n:

$$C_n^*=-\frac{1}{2}C_{1/2}^n(-4)^n=\frac{1}{n}C_{2(n-1)}^{n-1}=C_n.$$

($C_{1/2}^n$ 的定义可参阅本章第 2 节第 22 题.)

6. (二项系数 C_{N+n-1}^n 的性质与各种解释.) 证明,

(a) 由共有 $|A|=N$ 个元素的集合 A 中取出的大小为 n 的无序样本 $[\cdots]$(有重复元素, 即 "有放回抽样") 的数目等于 C_{N+n-1}^n.

(b) 由满足关系式 $n_1+\cdots+n_N=n$ 的非负整数 n_i, $i=1,\cdots,N$, 构成的有序数组的数目等于 C_{N+n-1}^n, $n\geqslant 1$, $N\geqslant 1$.

(c) n 个不可区分的小球放入 N 个不同的盒子 (每个盒子放入的球数不限, 即 "无限制的放球") 的放法数目等于 C_{N+n-1}^n, $n\geqslant 1$, $N\geqslant 1$.

提示: 参阅第 3 题的提示.

7. (参阅上题 (b).) 我们来观察方程 $n_1+\cdots+n_N=n\,(n\geqslant 1,\ N\geqslant 1)$ 的非负整数解 $n_i\geqslant 0$, $i=1,\cdots,N$ 的无序数组 $[n_1,\cdots,n_N]$. 这样的数组的个数有多少? 如果只限于正整数解 ($n_i>0$, $i=1,\cdots,N$), 那么相应的数组的个数有多少? 方程 $n_1+\cdots+n_N=n\,(n\geqslant 1,\ N\geqslant 1)$ 的正整数解的有序数组 $(n_1,\cdots,n_N)$ 的个数有多少?

8. (参阅第 6 题 (b) 和第 7 题.) 我们来考察关于非负整数或正整数 n_i, $i=1,\cdots,N$ 的不等式 $n_1+\cdots+n_N\leqslant n$. 试求满足不等式 $n_1+\cdots+n_N\leqslant n\,(n\geqslant 1,\ N\geqslant 1)$ 的有序解组 $(n_1,\cdots,n_N)$ 和无序解组 $[n_1,\cdots,n_N]$ 的个数.

9. 证明,

(a) n 条直线可以把平面 $\mathbb{R}^2$ 所分成的部分的最大数目为

$$1+\frac{n(n+1)}{2}.$$

(b) n $(n\geqslant 3)$ 个平面可以把空间 $\mathbb{R}^3$ 所分成的部分的最大数目为

$$\frac{1}{6}(n^3+5n+6).$$

10. 设 A 与 B 是 Ω 的两个子集. 证明, 由它们, (按照第 3 节的术语) 由子集类 $\mathscr{A}_0=\{A,B\}$, 所张成的代数 $\alpha(A,B)$ 共由 $N(2)=16$ 个元素组成:

$$\Big\{A,\ B,\ \overline{A},\ \overline{B},\ A\cap B,\ \overline{A}\cap\overline{B},\ A\setminus B,\ B\setminus A,$$
$$A\cup B,\ A\cup\overline{B},\ \overline{A}\cup B,\ \overline{A}\cup\overline{B},\ A\triangle B,\ \overline{A\triangle B},\ \Omega,\ \varnothing\Big\},$$

其中 $A\triangle B=(A\setminus B)\cup(B\setminus A)$ 为集合 A 与 B 的对称差 (参阅第二章第 1 节中的表格).

试给出 Ω 的分割 $\mathscr{D}$(参阅第 3 节), 使得由它所张成的代数 $\alpha(\mathscr{D})$ 与 $\alpha(A,B)$ 相同.

证明, 由子集类 $\mathscr{A}_0=\{A_1,\cdots,A_n\}$, 其中 $A_i\subseteq\Omega$, $i=1,\cdots,n$, 所张成的代数 $\alpha(A_1,\cdots,A_n)$ 由 $N(n)=2^{2^n}$ 个元素组成, 因而 $N(2)=16$, $N(3)=256$, 等等.

11. 证明布尔 (Boole) 不等式:

(a) $$\mathbf{P}\left(\bigcup_{i=1}^n A_i\right)\leqslant\sum_{i=1}^n\mathbf{P}(A_i),\quad \mathbf{P}\left(\bigcap_{i=1}^n A_i\right)\geqslant 1-\sum_{i=1}^n\mathbf{P}(\overline{A}_i).$$

证明, 对任何 $n\geqslant 1$, 都成立不等式

(b) $$\mathbf{P}\left(\bigcap_{i=1}^n A_i\right)\geqslant\sum_{i=1}^n\mathbf{P}(A_i)-(n-1),$$

且成立如下的古尼阿斯 (Кунилс) 不等式

(c) $$\mathbf{P}\left(\bigcup_{i=1}^n A_i\right)\leqslant\min_k\left\{\sum_{i=1}^n\mathbf{P}(A_i)-\sum_{i\neq k}\mathbf{P}(A_i\cap A_k)\right\}$$

和钟开莱–爱尔迪希 (Chung-Erdös) 不等式[①]

(d) $$\mathbf{P}\left(\bigcup_{i=1}^n A_i\right)\geqslant\frac{\left(\sum\limits_{i=1}^n\mathbf{P}(A_i)\right)^2}{\sum\limits_{i,j=1}^n\mathbf{P}(A_i\cap A_j)}.$$

①原文中的不等号错为 "$\leqslant$" —— 译者注.

提示: 对于上述不等式 (b) 在 $n=3$ 时的情形, 即 $\mathbf{P}(A_1\cap A_2\cap A_3)\geqslant\mathbf{P}(A_1)+\mathbf{P}(A_2)+\mathbf{P}(A_3)-2$, 可以直接通过观察来验证, 而对于其余情形, 则需采用归纳法.

12. 证明如下的关于事件 $A_1,\cdots,A_n\,(n\geqslant 1)$ 的并与交的概率容斥公式[①] (称为庞加莱 (Poincaré) 公式, *庞加莱定理*, *庞加莱恒等式*):

(a) $$\mathbf{P}(A_1\cup\cdots\cup A_n)=\sum_{1\leqslant i_1\leqslant n}\mathbf{P}(A_{i_1})-\sum_{1\leqslant i_1<i_2\leqslant n}\mathbf{P}(A_{i_1}\cap A_{i_2})+\cdots+$$
$$+(-1)^{m+1}\sum_{1\leqslant i_1<\cdots<i_m\leqslant n}\mathbf{P}(A_{i_1}\cap\cdots\cap A_{i_m})+\cdots+(-1)^{n+1}\mathbf{P}(A_1\cap\cdots\cap A_n)$$

和

(b) $$\mathbf{P}(A_1\cap\cdots\cap A_n)=\sum_{1\leqslant i_1\leqslant n}\mathbf{P}(A_{i_1})-\sum_{1\leqslant i_1<i_2\leqslant n}\mathbf{P}(A_{i_1}\cup A_{i_2})+\cdots+$$
$$+(-1)^{m+1}\sum_{1\leqslant i_1<\cdots<i_m\leqslant n}\mathbf{P}(A_{i_1}\cup\cdots\cup A_{i_m})+\cdots+(-1)^{n+1}\mathbf{P}(A_1\cup\cdots\cup A_n).$$

注 1. 公式 (a) 和 (b) 可以缩写为下述形式:

$$\mathbf{P}\left(\bigcup_{i=1}^n A_i\right)=\sum_{m=1}^n(-1)^{m+1}S_m,\quad \mathbf{P}\left(\bigcap_{i=1}^n A_i\right)=\sum_{m=1}^n(-1)^{m+1}\widetilde{S}_m,$$

其中

$$S_m=\sum_{1\leqslant i_1<\cdots<i_m\leqslant n}\mathbf{P}(A_{i_1}\cap\cdots\cap A_{i_m}),\quad \widetilde{S}_m=\sum_{1\leqslant i_1<\cdots<i_m\leqslant n}\mathbf{P}(A_{i_1}\cup\cdots\cup A_{i_m}).$$

注 2. 虽然介绍容斥公式是《概率》(第一卷) 第一章中的内容, 在那里仅仅涉及了有限的概率空间 $(\Omega,\mathscr{A},\mathbf{P})$, 我们需要强调的是, 这个公式对于一般的概率空间 $(\Omega,\mathscr{F},\mathbf{P})$ 也是适用的 (参阅第二章).

注 3. 我们用 $|A|$ 表示有限集合 Ω 的子集 A 中的元素个数. 那么在 $\mathbf{P}(A)=|A|/|\Omega|$(古典概型) 的场合下, 可以由公式 (a) 和 (b) 推出关于有限集合 $A_1,\cdots,A_n$ 的容斥公式:

(a′) $$\left|\bigcup_{i\in T}A_i\right|=\sum_{\varnothing\neq S\subseteq T}(-1)^{|S|+1}\left|\bigcap_{i\in S}A_i\right|;$$

(b′) $$\left|\bigcap_{i\in T}A_i\right|=\sum_{\varnothing\neq S\subseteq T}(-1)^{|S|+1}\left|\bigcup_{i\in S}A_i\right|,$$

其中 $T=\{1,\cdots,n\}$.

①也称为进出公式 —— 译者注.

提示: 公式 (a) 可以运用归纳法来证明, 首先应当证明 $n=2$ 时的公式

$$\mathbf{P}(A_1 \cup A_2) = [\mathbf{P}(A_1) + \mathbf{P}(A_2)] - \mathbf{P}(A_1 \cap A_2)$$

成立 (亦可参阅第 4 节第 9 题).

为了证明公式 (b), 需要注意

$$\mathbf{P}\left(\bigcap_{i=1}^n A_i\right) = \mathbf{P}\left(\overline{\bigcup_{i=1}^n \overline{A}_i}\right) = 1 - \mathbf{P}\left(\bigcup_{i=1}^n \overline{A}_i\right),$$

然后对 $\overline{A}_1, \cdots, \overline{A}_n$ 运用公式 (a).

13. 用 B_m 表示事件 $A_1, \cdots, A_n$ 中恰有 m 个同时发生的事件, $0 \leqslant m \leqslant n$. 证明,

$$\mathbf{P}(B_m) = \sum_{k=m}^n (-1)^{k-m} C_k^m S_k,$$

或者, 展开为

$$\mathbf{P}(B_m) = S_m - C_{m+1}^1 S_{m+1} + \cdots + (-1)^{n-m} C_n^{n-m} S_n.$$

再用 $B_{\geqslant m}$ 表示事件 $A_1, \cdots, A_n$ 中至少有 m 个同时发生的事件, 试由上面的公式推出

$$\mathbf{P}(B_{\geqslant m}) (= \mathbf{P}(B_m) + \cdots + \mathbf{P}(B_n)) = \sum_{k=m}^n (-1)^{k-m} C_{k-1}^{m-1} S_k,$$

或者, 等价地写为

$$\mathbf{P}(B_{\geqslant m}) = S_m - C_m^1 S_{m+1} + \cdots + (-1)^{n-m} C_{n-1}^{n-m} S_n.$$

提示: 关于 $\mathbf{P}(B_m)$ 的公式 (称为华林 (Waring) 公式) 可以运用第 12 题中的容斥公式给出一个组合证明 (在费勒 (Feller) 的书中有一个这样的证明 (参考文献 [119], 第 1 卷第 4 章第 3 节)). 建议那些已经掌握了随机变量和数学期望概念的读者, 采用下述证明方法:

令 $X_i = I_{A_i}$, 即事件 A_i 的示性函数, $i = 1, \cdots, n$. 假设对于给定的试验结果 ω, 标号为 $i_1, \cdots, i_m$ 的事件发生, 而标号为 $j_1, \cdots, j_{n-m}$ 的事件没有发生.

我们来观察和数

$$\sum X_{i_1} \cdots X_{i_m} (1 - X_{j_1}) \cdots (1 - X_{j_{n-m}}),$$

其中的求和对一切可能的选取 $(i_1, \cdots, i_m)$ 和 $(j_1, \cdots, j_{n-m})$ 进行. 如果试验结果 ω 恰好属于 m 个事件 $A_{i_1}, \cdots, A_{i_m}$, 那么该和等于 m; 在其余情形下, 该和为 0. 而概率 $\mathbf{P}(B_m)$ 就是所求的和的数学期望. 接下来的步骤可以参阅第 4 节第 9 题的提示 (也可参阅第二章第 6 节第 31 题).

14. 试由第 13 题中关于 $\mathbf{P}(B_m)$ 和 $\mathbf{P}(B_{\geqslant m})$ 的公式推出下述对一切偶数 r 成立的邦费尔罗尼 (Bonferroni) 公式:

$$S_m + \sum_{k=1}^{r+1} (-1)^k C_{m+k}^k S_{m+k} \leqslant \mathbf{P}(B_m) \leqslant S_m + \sum_{k=1}^{r} (-1)^k C_{m+k}^k S_{m+k},$$

$$S_m + \sum_{k=1}^{r+1} (-1)^k C_{m+k-1}^k S_{m+k} \leqslant \mathbf{P}(B_{\geqslant m}) \leqslant S_m + \sum_{k=1}^{r} (-1)^k C_{m+k-1}^k S_{m+k}.$$

提示: 本题的证法之一是首先验证下述 (有益的) 表达式的正确性:

$$S_m = \sum_{r=m}^n C_r^m \mathbf{P}(B_m), \quad S_m = \sum_{r=m}^n C_{r-1}^{m-1} \mathbf{P}(B_{\geqslant m}).$$

15. 利用第 12 题中关于 S_m 和 $\widetilde{S}_m$ 的定义, 证明

(a) 邦费尔罗尼不等式 (上题中不等式的特殊情形):

$$S_1 - S_2 + \cdots - S_{2k} \leqslant \mathbf{P}\left(\bigcup_{i=1}^n A_i\right) \leqslant S_1 - S_2 + \cdots + S_{2k-1},$$

其中 $k \geqslant 1$, 使得 $2k \leqslant n$.

(b) 古特别里不等式:

$$\mathbf{P}\left(\bigcup_{i=1}^n A_i\right) \leqslant \frac{\widetilde{S}_m}{C_{n-1}^{m-1}}, \quad m = 1, \cdots, n$$

($m=1$ 时就是第 11 题布尔不等式中的第一个不等式).

(c) 弗雷歇 (Fréchet) 不等式:

$$\mathbf{P}\left(\bigcup_{i=1}^n A_i\right) \leqslant \frac{\widetilde{S}_{m+1}}{C_{n-1}^m} \leqslant \frac{\widetilde{S}_m}{C_{n-1}^{m-1}}, \quad m = 0, 1, \cdots, n-1.$$

16. (无序问题) 设 $(i_1, \cdots, i_n)$ 为正整数 $1, \cdots, n$ 的随机排列 (各种排列方式的发生概率都是 $1/n!$). 证明

(a) 在 $1, \cdots, n$ 的排列中恰有 m 个数位于自己的位置 (部分无序) 上的概率 $P_{(m)}$ 等于

$$\frac{1}{m!}\left(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^{n-m} \frac{1}{(n-m)!}\right) \quad \left(\sim \frac{e^{-1}}{m!},\ n \to \infty\right).$$

(b) 在 $1, \cdots, n$ 的排列中至少有 1 个数位于自己的位置上的概率 $P_{(\geqslant 1)}$ 等于

$$1 - \frac{1}{2!} + \frac{1}{3!} + \cdots + (-1)^{n-1} \frac{1}{n!} \quad (\sim 1 - e^{-1},\ n \to \infty),$$

由此可知, 完全无序 (即任何一个数都不在自己的位置上) 的概率等于

$$1-P_{(\geqslant 1)}=\sum_{j=0}^{n}\frac{(-1)^j}{j!}\quad (\sim e^{-1},\quad n\to\infty).$$

提示: 设 A_i 表示集合 $\{1,\cdots,n\}$ 中的数 i 在排列中确实位于自己的位置上的事件. 那么非常有趣的是: 此处的概率 $P_{(m)}$ 就等于第 13 题中的概率 $\mathbf{P}(B_m)$, 亦即

$$P_{(m)}=S_m-C_{m+1}^{1}S_{m+1}+\cdots+(-1)^{n-m}C_n^{n-m}S_n.$$

事实上, 只要证明了 $S_k=1/k!$, $m\leqslant k\leqslant n$, 就可得到该公式.

为了证明关于 $P_{(\geqslant 1)}$ 的公式, 仍然需要利用第 13 题, 即

$$P_{(\geqslant 1)}=\mathbf{P}(B_{\geqslant 1})=S_1-S_2+S_3-\cdots+(-1)^{n-1}S_n,$$

其中 $S_k=\frac{(n-k)!}{n!}$.

17. (重合问题) 设有 n 封信和 n 个信封. 信被随机地装入各个信封, 亦即对各种不同的装入方式按照 "古典概型" 赋概 (参阅第 5 节). 以 $P_{(m)}$ 表示恰有 m 封信被装入自己的信封的概率.

证明

$$P_{(m)}=\frac{1}{m!}\sum_{j=0}^{n-m}\frac{(-1)^j}{j!}.\text{①}$$

提示: 1) 自然首先应当考虑如何理解 "信被随机地装入信封" 的提法. 如果将其理解为: 1 号信可以随意放入 n 个信封中的任何一个, 2 号信可以随意放入剩下的 $n-1$ 个信封中的任何一个, 如此一直下去, 那么这样的过程就可以看成为一种 "无放回抽样", 与信的号码 $(1,\cdots,n)$ 相对应的是信封号码的有序数组 $(a_1,\cdots,a_n)$, 因此共有 $(n)_n=n!$ 种不同的放入方式. 再根据经典的 "等可能性" 原则 (参阅第 5 节), 得到每一种顺序 $(a_1,\cdots,a_n)$ 的概率都是 $1/n!$.

2) 用 A_i 表示第 i 号信被放入 "自己的" 第 i 号信封的事件, 那么就有 $P_{(m)}=\mathbf{P}(B_m)$(参阅第 13 题), 并且

$$P_{(m)}=\sum_{k=m}^{n}(-1)^{k-m}C_k^m S_k.$$

于是只要证得 $S_k=1/k!$, $1\leqslant k\leqslant n$, 即可得到所要证明的关于 $P_{(m)}$ 的公式. (我们指出, $P_{(0)}$ 表示任何一封信都没有被放入自己信封的概率, 其值为 $\sum\limits_{k=2}^{n}(-1)^k\frac{1}{k!}$, 甚至对于固定的 n $(n>4)$, 该值都非常接近于 e^{-1}.)②

①原书为 $P_{(m)}=\frac{1}{m!}\left(1-\sum\limits_{j=0}^{n-m}\frac{(-1)^j}{j!}\right)$ —— 译者注.

②原书为 $1-e^{-1}$ —— 译者注.

18. 从幼儿园离开的时候, n 个孩子中的每一人都以 "随机方式" 取了一只左脚靴子和一只右脚靴子. 证明,

(a) 若以 P_a 表示他们全都没有穿着自己的一双靴子离开的概率, 则有

$$P_a=\sum_{k=0}^{n}(-1)^k\frac{(n-k)!}{k!\,n!}.\text{①}$$

(b) 若以 P_b 表示他们全都既没有穿着自己的左脚靴子, 又没有穿着自己的右脚靴子离开的概率, 则有

$$P_b=\left(\sum_{k=2}^{n}(-1)^k\frac{1}{k!}\right)^2.$$

提示: 首先应当明确 "以随机方式" 的含义. (可以按照第 17 题提示中的方式类似地解决).

(a) 若以 A_i 表示第 i 个孩子取了自己的左脚靴子和自己的右脚靴子的事件. 则由容斥公式可得

$$P_a=\mathbf{P}\left(\bigcap_{i=1}^{n}\overline{A}_i\right)=1-\mathbf{P}\left(\bigcup_{i=1}^{n}A_i\right)=1-S_1+S_2-\cdots+(-1)^nS_n,$$

应当指出, 此处与题目中关于 P_a 的表达式相同, 有 $S_k=\frac{(n-k)!}{k!n!}$.

(b) 为了证明关于 P_b 的表达式, 应当明确

$$P_b=\mathbf{P}(\text{每个孩子都没有穿着自己的左脚靴子})\cdot\mathbf{P}(\text{每个孩子都没有穿着自己的右脚靴子}),$$

再利用第 17 题中的结果.

19. 我们来考察将 n 个质点按麦克斯韦–波尔查诺统计方式放入 M 个盒子 ("不同的质点, 不同的盒子, 无限制地放入") 的问题. 按照拉普拉斯古典赋概方法, 利用第 1 节中公式 (10)(与摆放质点的 "随机方式" 相对应), 证明, 在指定的盒子中恰放入 k 个质点的概率 $P_k(n;M)$ 满足如下公式:

$$P_k(n;M)=C_n^k\frac{(M-1)^{n-k}}{M^n}.$$

并由此推出: 当 $n\to\infty$, $M\to\infty$, 并且 $n/M\to\lambda>0$ 时, 有

$$P_k(n;M)\to e^{-\lambda}\frac{\lambda^k}{k!}.$$

(试与泊松分布相比较, 参阅第 6 节和第二章第 3 节中的表 2.)

①原书为 $P_a=\sum\limits_{k=1}^{n}(-1)^k\frac{(n-k)!}{k!\,n!}$ —— 译者注.

20. (第 19 题续) 以 $R_m(n;M)$ 表示恰有 m 个盒子空着的概率. 证明

$$R_m(n;M) = C_M^m \sum_{k=0}^{M-m} (-1)^k C_{M-m}^k \left(1 - \frac{m+k}{M}\right)^n.$$

并由此证明, 如果 $n \to \infty$, $M \to \infty$, 并且 $Me^{-n/M} \to \lambda > 0$, 则有

$$R_m(n;M) \to e^{-\lambda}\frac{\lambda^m}{m!}.$$

21. 仍然考察将 n 个质点放入 M 个盒子的问题, 不过这次按照玻色–爱因斯坦统计方式进行 ("无区别的质点, 不同的盒子, 无限制地放入"). 以 $Q_k(n;M)$ 表示在指定的盒子中恰放入 k 个质点的概率.

证明,

$$Q_k(n;M) = \frac{C_{M+n-k-2}^{n-k}}{C_{M+n-1}^n},$$

又当 $n \to \infty$, $M \to \infty$, 并且 $n/M \to \lambda > 0$ 时, 有

$$Q_k(n;M) \to p(1-p)^k, \quad 其中\ p = \frac{1}{1+\lambda}.$$

(试与几何分布相比较, 参阅第二章第 3 节中的表 2.)

22. 罐中装有分别写有编号 $1, \cdots, M$ 的 M 个小球, 从中有放回地随机取出 n 个小球. 以 A_k 表示所取出的小球的最大号码是 k 的事件. 证明

$$\mathbf{P}(A_k) = \frac{k^n - (k-1)^n}{M^n}.$$

并且证明, 如果采用无放回的取球方式, 则有

$$\mathbf{P}(A_k) = \frac{C_{k-1}^{n-1}}{C_M^n}, \quad n \leqslant k \leqslant M.$$

23. 证明关于两个函数 f 与 g 乘积求导的莱布尼茨公式:

$$D^N(fg) = \sum_{n=0}^{N} C_N^n (D^n f)(D^{N-n} g).$$

提示: 利用归纳法和帕斯卡三角形的性质, 即 $C_{N+1}^n = C_N^n + C_N^{n-1}$(参阅第 2 节第 2 题).

§2. 某些经典模型和分布

1. 证明,

$$(x+y)^n = \sum_{k=0}^{n} C_n^k x^k y^{n-k} \quad (二项恒等式),$$

$$(x+y)_n = \sum_{k=0}^{n} C_n^k (x)_k (y)_{n-k} \quad (范德蒙德恒等式),$$

$$[x+y]_n = \sum_{k=0}^{n} C_n^k [x]_k [y]_{n-k} \quad (廖尔伦德恒等式),$$

其中

$$(x)_n = x(x-1)(x-2)\cdots(x-n+1),$$
$$[x]_n = x(x+1)(x+2)\cdots(x+n-1).$$

提示: 利用关于多项式的泰勒公式.

2. 试运用各种观察方法, 例如概率观察、组合观察 (计算有利场合下的方法数目)、几何观察 (计算由一个地点到另一个地点的路径条数) 以及其他手段 (例如, 代数方法, 包括计算与比较形如 $(1+x)^a(1+x)^b = (1+x)^{a+b}$ 的恒等式两端的 x^n 的系数, 等等), 证明组合系数的下述性质:

$$1 = C_n^0 < C_n^1 < \cdots < C_n^{\lfloor n/2 \rfloor} = C_n^{\lceil n/2 \rceil} > \cdots > C_n^{n-1} > C_n^n = 1,$$

即对称性与单峰性 ($\lfloor x \rfloor \equiv [x]$ 表示不超过实数 x 的最大整数, 称为 x 的地板函数或高斯函数, $\lceil x \rceil$ 表示不小于实数 x 的最小整数, 称为 x 的天花板函数);

$C_N^{k-1} + C_N^k = C_{N+1}^k$, 即帕斯卡三角形性质:

```
                  1
                1   1
              1   2   1
            1   3   3   1
          1   4   6   4   1
        1   5   10  10  5   1
      1   6   15  20  15  6   1
      ...........................
```

$$C_n^k = C_{n-2}^k + 2C_{n-2}^{k-1} + C_{n-2}^{k-2},$$

$$\sum_{k=0}^{N} C_N^k = 2^N, \quad \sum_{k=0}^{N} (C_N^k)^2 = C_{2N}^N, \quad \sum_{k=0}^{N} 2^k C_N^k = 3^N, \quad \sum_{k=1}^{N} k C_N^k = N \cdot 2^{N-1},$$

$$\sum_{k=0}^{N} (-1)^{N-k} C_M^k = C_{M-1}^N, \quad M \geqslant N+1,$$

$$\sum_{k=0}^{N} k(k-1) C_N^k = N(N-1)2^{N-2}, \quad N \geqslant 2,$$

$$kC_N^k = NC_{N-1}^{k-1}, \quad \sum_{m=k}^{N} C_m^k = C_{N+1}^{k+1}, \quad C_N^k C_k^l = C_N^l C_{N-l}^{k-l}, \quad l \leqslant k \leqslant N,$$

$$\sum_{j=0}^{k} C_N^j = \sum_{j=0}^{k} 2^j C_{N-1-j}^{k-j}, \quad k \leqslant N-1, \quad \sum_{j=0}^{k} C_{N+j}^j = C_{N+k+1}^k,$$

$$C_{N-k}^k + C_{N-k-1}^{k-1} = \frac{N}{N-k} C_{N-k}^k, \quad 0 \leqslant k \leqslant N,$$

$$C_{N_1+N_2}^n = \sum_{k=0}^{N_1} C_{N_1}^k C_{N_2}^{n-k} \quad (\text{范德蒙德二项卷积}),$$

$$C_N^{n-1} \leqslant C_N^n, \quad 1 \leqslant n \leqslant \frac{N+1}{2}, \quad C_N^{n-1} C_N^{n+1} \leqslant (C_N^n)^2, \quad n \leqslant N-1,$$

$$C_{M+N}^N \leqslant \left(1+\frac{M}{N}\right)^N \left(1+\frac{N}{M}\right)^M \text{ 或者, 等价地 } \frac{(M+N)!}{(M+N)^{M+N}} \leqslant \frac{M!N!}{M^M N^N},$$

$$\frac{N^k}{k^k} \leqslant C_N^k \leqslant \frac{N^k}{k!}, \quad \sum_{k=0}^{N} C_N^k (-1)^{N-k} 2^k = 1,$$

$$C_N^n = \sum_{k=n}^{N} (-1)^{N+k} C_N^k C_k^n 2^{k-n}, \quad 1 \leqslant n \leqslant N,$$

$$\sum_{k=0}^{N} C_N^k C_k^n = 2^{N-n} C_N^n, \quad 1 \leqslant n \leqslant N,$$

$$\sum_{k=0}^{N} (-1)^k (C_N^k)^2 = \begin{cases} (-1)^m C_{2m}^m, & \text{如果 } N = 2m, \\ 0, & \text{其他情形}, \end{cases}$$

$$\sum_{k=0}^{N} (-1)^k (C_N^k)^3 = \begin{cases} (-1)^m (3m)!(m!)^{-3}, & \text{如果 } N = 2m, \\ 0, & \text{其他情形}, \end{cases}$$

$$\sum_{k=1}^{N} C_N^k \frac{(-1)^{k-1}}{k} = \sum_{k=1}^{N} \frac{1}{k},$$

$$\sum_{k=0}^{N} (-1)^{N-k} k^l C_N^k = \begin{cases} 0, & \text{如果 } l < N, \\ N!, & \text{如果 } l = N, \end{cases}$$

$$\sum_{k=0}^{N} C^{N-k} \cdots C^{k} \cdots 2^{2N}$$

(亦可参阅第 22 题.)

3. 设 p 为素数, $1 \leqslant k \leqslant p-2$, $p \geqslant 3$. 证明, p 可整除 C_p^k, 且 $C_{2p}^p = 2 \pmod p$.

4. 证明, 将正整数 N 自身拆为两个部分 (即将正整数 N 自身表示为两个非负整数的和) 的方法有 $[N/2]+1$ 种, 其中 $[x]$ 是实数 x 的整数部分.

5. 正如本书最后的附录所言, 二阶斯特林数 S_N^n 表示将一个 N 元集合, 例如 $\{1, \cdots, N\}$, 拆分为 n 个非空子集的方法数目 (见附录, 第 312 页).

证明如下性质成立:

(a) $S_N^1 = S_N^N = 1, \quad S_N^2 = 2^{N-1}-1, \quad S_N^{N-1} = C_N^2.$

(b) $S_{N+1}^n = S_N^{n-1} + nS_N^n, \quad 1 \leqslant n \leqslant N.$

(c) $S_{N+1}^n = \sum_{k=0}^{N} C_N^k S_k^{n-1}$ (当$p > k$时, $S_k^p = 0$).

(d) $S_N^n = \frac{1}{n!} \sum_{k=0}^{N-1} (-1)^k C_n^k (n-k)^N.$

(e) $S_N^n = \frac{1}{n!} \sum_{k=0}^{n} (-1)^{n-k} C_n^k k^N.$

(f) $S_N^n < n^{N-n} C_{N-1}^{n-1}.$

提示: 证明性质 (b) 是证明接下来的性质 (c) 和 (d) 的关键, 而为了证明性质 (b), 需要利用关系式 $x^N = \sum_{k=0}^{N} S_N^n (x)_n$(参阅附录, 第 323 页) 和 $(x)_{n+1} = (x-n)(x)_n$.

6. 在附录的第 3 节中, (利用关系式 $x^N = \sum_{k=0}^{N} S_N^n (x)_n$) 证明了: 二阶斯特林数序列 $S^n = (S_N^n)_{n \geqslant 0}$ 的指数型母函数

$$E_{S^n}(x) = \sum_{N \geqslant 0} S_N^n \frac{x^N}{N!}$$

由如下关系式给出:

$$E_{S^n}(x) = \frac{(e^x-1)^n}{n!}.$$

试运用上题中的性质 (e) 重新导出这一关系式.

7. 根据二阶斯特林数的定义方式之一 (参阅附录第 3 节, 第 324 页), $(-1)^{N-n} s_N^n$ 是正整数 $\{1, \cdots, N\}$ 的恰有 n 个循环圈的排列的数目 ($s_N^0 = 0$).

证明,

(a) $s_N^N = 1, \quad s_N^1 = (N-1)!(-1)^{N-1}$.

(b) $s_{N+1}^n = s_N^{n-1} - N s_N^n, \quad 1 \leqslant n \leqslant N$.

(c) $\sum_{n=1}^{N}(-1)^{N-n} s_N^n = \sum_{n=1}^{N} |s_N^n| = N!$.

并且证明 s_N^n, $1 \leqslant n \leqslant N$, 满足代数关系式 (亦可参阅附录第 324 页)

(d) $(x)_N = \sum_{n=0}^{N} s_N^n x^n$,

其中 $(x)_N = x(x-1)\cdots(x-N+1)$.

提示: 递推关系式 (b) 既可用组合方法证明, 也可直接由代数关系式 (d) 推出.

8. 证明, 一阶和二阶斯特林数具有如下的二重性:

$$\sum_{n \geqslant 0} S_N^n s_n^M = \delta_{NM},$$

其中 δ_{ab} 为克罗内克符号($\delta_{aa} = 1$, 如果 $a \neq b$, 则 $\delta_{ab} = 0$).

9. 证明, 一阶斯特林数序列 $s^n = (s_N^n)_{n \geqslant 0}$ 的指数型母函数

$$E_{s^n}(x) = \sum_{N \geqslant 0} s_N^n \frac{x^N}{N!}$$

由如下关系式所定义:

$$E_{s^n}(x) = \frac{(\ln(1+x))^n}{n!}.$$

10. 根据定义 (参阅附录第 312 页), 贝尔数 B_N 表示集合 $\{1, 2, \cdots, N\}$ $(N \geqslant 1)$ 的所有可能的拆分方式数目, 亦即

$$B_N = \sum_{n=1}^{N} S_N^n,$$

其中 S_N^n 是二阶斯特林数.

令 $B_0 = 1$, 证明,

(a) 递推关系式

$$B_N = \sum_{k=1}^{N} C_{N-1}^{k-1} B_{N-k}.$$

(b) 指数型母函数 $E_B(x) = \sum_{N \geqslant 0} B_N \frac{x^N}{N!}$ 由如下关系式给出:

$$E_B(x) = \exp\{e^x - 1\}.$$

(c) $B_N < N!$ 且 $\lim_{N \to \infty} \left(\frac{B_N}{N!}\right)^{1/N} = 0$.

验证 $B_1, \cdots, B_5$ 分别等于 $1, 2, 5, 15, 52$.

提示: 在证明 (b) 时, 需要利用 (a), 首先证明 $E_B(x)$ 满足微分方程

$$\frac{\mathrm{d}E_B(x)}{\mathrm{d}x} = e^x E_B(x)$$

且初值为 $E_B(0) = 1$. 为证 (c) 中的极限式, 需要考察级数 $\sum_{N \geqslant 0} B_N \frac{x^N}{N!}$ 的收敛半径 $R = 1/\overline{\lim}\left(\frac{B_N}{N!}\right)^{1/N}$(不难看出, 该级数在实直线上对任何 x 都收敛).

11. (斐波那契数) 设 $n \geqslant 1$, 用 F_n 记将正整数 n 表示为若干个 1 和 2 的有序和的不同方式数目. 易知 $F_1 = 1, F_2 = 2$(因为 $2 = 1+1 = 2$), $F_3 = 3$(3=1+1+1=1+2=2+1), $F_4 = 5$(4=1+1+1+1=2+1+1=1+2+1=1+1+2=2+2).

(a) 证明, 等于 $n \geqslant 2$, 斐波那契数 F_n 满足递推关系式

$$F_n = F_{n-1} + F_{n-2}, \qquad (*)$$

其中 $F_0 = 1, F_1 = 1$.

(b) 由递推关系式 (*) 推出

$$F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1}\right], \qquad (**)$$

其中 $\frac{1+\sqrt{5}}{2} = 1.6180\cdots$, $\frac{1-\sqrt{5}}{2} = -0.6180\cdots$.

(c) 利用递推关系式 (*) 证明, 序列 $(F_n)_{n \geqslant 0}$ 的母函数 $F(x) = \sum_{n \geqslant 0} F_n x^n$ 由下式给出

$$F(x) = \frac{1}{1 - x - x^2}. \qquad (***)$$

(d) 斐波那契数 (最初见 18 世纪初斐波那契的 "算盘学" 讲义中关于克罗内克分解问题的习题) 有着许多有趣的性质, 例如

$$F_1 + F_2 + \cdots + F_n = F_{n+1} - 1, \quad F_{n-1}^2 + F_n^2 = F_{2n},$$
$$F_{n-1}F_n + F_nF_{n+1} = F_{2n+1}, \quad F_mF_n + F_{m-1}F_{n-1} = F_{m+n},$$

其中 $m, n \geqslant 0$, 而 $F_{-1} = 0$.

试验证上列等式.

(e) 证明, 斐波那契数满足关系式

$$F_n = \sum_{k=0}^{[n/2]} C_{n-k}^k.$$

(可以验证 $(F_0, F_1, \cdots, F_{17}) = (1,1,2,3,5,8,13,21,34,55,89,144,233,377,610,987,1597,2584)$.)

(f) 证明, 对于 $n \leqslant 9$, 有

$$\frac{F_n}{\lceil e^{(n-1)/2} \rceil} = 1,$$

而 (对于 $n=10$) $F_{10}/\lceil e^{9/2} \rceil < 1$ 并且有

$$\lim_{n\to\infty} \frac{F_n}{\lceil e^{(n-1)/2} \rceil} = 0. \qquad (****)$$

提示: 为证 (b) 中的 (**) 式, 首先应当找出递推式 $F_n = F_{n-1} + F_{n-2}$ 的形如 $F_n = \alpha^n$ 的解. 还可通过考察函数 (***) 的泰勒展开式中的 x^n 的系数来证明 (**) 式. (应当指出: $1 - x - x^2 = (1 - ax)(1 - bx)$, 其中 $a = (1+\sqrt{5})/2$, $b = (1-\sqrt{5})/2$.)

在证明 (f) 中的 (****) 式时, 应当注意 $F_n \sim c_1 \cdot (1.618\cdots)^n$, 而 $\lceil e^{(n-1)/2} \rceil \sim c_2 \cdot (1.648\cdots)^n$, 而 c_1 和 c_2 皆为常数 (试求出它们).

12. 证明, 对于多项系数 (多项式系数)

$$C_N(n_1, \cdots, n_r) = \frac{N!}{n_1! \cdots n_r!}, \quad n_1 + \cdots + n_r = N, \quad n_i \geqslant 0$$

成立如下的关系式 (范德蒙德多项卷积):

$$C_{N_1+N_2}(n_1, \cdots, n_r) = \sum C_{N_1}(k_1, \cdots, k_r) C_{N_2}(n_1 - k_1, \cdots, n_r - k_r),$$

其中求和对所有满足如下条件的整数 $k_i \geqslant 0$ 进行: $k_i \leqslant n_i$, $\sum_{i=1}^{r} k_i = N_1$, $\sum_{i=1}^{r} n_i = N_1 + N_2$.

13. 证明

$$(x_1 + \cdots + x_r)^N = \sum C_N(n_1, \cdots, n_r) x_1^{n_1} \cdots x_r^{n_r},$$

其中求和对所有满足条件 $\sum_{i=1}^{r} n_i = N$ 的整数 $n_i \geqslant 0$ 进行.

14. 证明, 非负格点集合 $\mathbb{Z}_+^r = \{(i_1, \cdots, i_r): \ i_1, \cdots, i_r = 0, 1, 2, \cdots\}$ 上的以原点 $(0, \cdots, 0)$ 为起点, 以 $(n_1, \cdots, n_r)$ 为终点 (其中 $\sum_{i=1}^{r} n_i = N$) 的非降路径的条数等于 $C_N(n_1, \cdots, n_r)$.(所谓非降路径, 就是每一步都只有一个坐标值增加一个单位的路径.)

15. 设 A 与 B 均为由有限个元素构成的集合 (分别有 $|A|$ 和 $|B|$ 个元素).

对于函数 $\mathbb{F}: A \to B$, 我们理解为对于每个 $a \in A$, 都有某个 $b \in B$ 与之对应的法则.

对于单射 $\mathbb{I}: A \to B$, 我们理解为不同的 $a \in A$, 对应为不同的 $b \in B$ 的法则 (此时有 $|A| \leqslant |B|$).

对于满射 $\mathbb{S}: A \to B$, 我们理解为对于每个 $b \in B$, 都能找到某个 $a \in A$, 使得 $b = \mathbb{S}(a)$ 的对应法则 (此时有 $|A| \geqslant |B|$).

对于双射 $\mathbb{B}: A \to B$, 我们理解为既是单射, 又是满射的对应法则 (此时有 $|A| = |B|$).

如果 $|A| = N, |B| = M$, 证明, 不同的函数、单射、满射和双射的对应法则的数目分别为

$$N(\mathbb{F}) = M^N, \quad N(\mathbb{I}) = (M)_N, \quad N(\mathbb{S}) = M! S_N^M, \quad N(\mathbb{B}) = N!.$$

16. 设 $P_N = \sum_{n=0}^{N} (N)_n$ 为从 N 个不同元素中取出若干个元素进行排列的方式数目的总和 ($(N)_0 = 1$, $(N)_n = N(N-1)(N-2)\cdots(N-n+1)$), $N \geqslant 1$. 证明, P_N 满足如下的递推关系式 ($P_0 = 1$):

$$P_N = NP_{N-1} + 1, \quad N \geqslant 1.$$

证明

$$P_N = N! \sum_{n=0}^{N} \frac{1}{n!},$$

并且证明 P_N 是与 $eN!$ 最接近的整数.

17. 证明, 序列 $P = (P_N)_{N \geqslant 0}$ (其中 P_N ($N \geqslant 0$) 的定义见上题) 的指数型母函数 $E_P(x) = \sum_{N=0}^{\infty} P_N \frac{x^N}{N!}$ 由如下关系式给出:

$$E_P(x) = \frac{e^x}{1-x}.$$

18. 证明 (5) 式.[①]

提示: 在所指的取极限过程中, 应当利用

$$\lim \mathbf{P}(B_{n_1,n_2}) = \lim \frac{C_n^{n_1} M_1^{n_1} M_2^{n_2}}{M^n}.$$

19. 证明, 多项分布 $\{\mathbf{P}(A_{n_1,\cdots,n_r})\}$ 的最大概率值在满足不等式 $np_i - 1 < k_i \leqslant (n+r-1)p_i$ $(i = 1, \cdots, r)$ 的点 $(k_1, \cdots, k_r)$ 处达到.

---

①此处所要证明的 (5) 式为: 设 n 与 M 为正整数, n_1, n_2, M_1, M_2 为非负整数, 有 $n_1 + n_2 = n$, $M_1 + M_2 = M$, 而

$$\mathbf{P}(B_{n_1,n_2}) = \frac{C_{M_1}^{n_1} C_{M_2}^{n_2}}{C_M^n}.$$

证明, 如果 $M \to \infty$, $M_1 \to \infty$, 并且 $\frac{M_1}{M} \to p$, 则有

$$\mathbf{P}(B_{n_1,n_2}) \to C_{n_1+n_2}^{n_2} p^{n_1} (1-p)^{n_2}$$

—— 译者注.

20. (一维伊金格 (Ezenger) 模型) 设有 n 个质点, 分别位于点 $1, \cdots, n$ 上.

假设每个质点都属于两种类型之一, 其中第一种类型的质点有 n_1 个, 第二种类型的质点有 n_2 个 $(n_1+n_2=n)$. 并且假定所有 $n!$ 种排列是等可能的.

试构作一个概率模型, 并求出事件 $A_n(m_{11}, m_{12}, m_{21}, m_{22}) = \{\nu_{11}=m_{11}, \nu_{12}=m_{12}, \nu_{21}=m_{21}, \nu_{22}=m_{22}\}$ 的概率, 其中 ν_{ij} 表示位于第 j 种类型的质点后面的第 i 种类型的质点的数目, $i,j=1,2$.

21. 设 N 为某个群体的大小, 要求不通过逐个数数的途径, 而是利用 "尽可能少的" 资料估计出 N 来. 详说这个问题十分有趣, 例如, 估计某个国家, 或是某个城市的人口数量, 等等.

1786 年, 拉普拉斯为估计法国的人口数量, 提出了如下的办法.

从群体中取出部分个体, 例如 M 个个体, 把它们做上记号后放回群体. 假定它们已经与其他个体 "很好地混合" 在一起. 此后再从混合后的群体中取出 n 个个体, 以 X 表示其中 (这 n 个个体中) 有记号的个体数目.

(a) 证明, $X=m$(在固定的 N, M, n 之下) 的概率 $\mathbf{P}_{N,M;n}\{X=m\}$ 由超几何分布公式给出:

$$\mathbf{P}_{N,M;n}\{X=m\} = \frac{C_M^m C_{N-M}^{n-m}}{C_N^n}.$$

(b) 假定 M, n, m 给定, 寻找使得 $\mathbf{P}_{N,M;n}\{X=m\}$ 达到最大的 N, 亦即寻求导致有记号的个体为 m 的总体的 "极大似然" 的大小.

证明, 这样找出的极大似然值 $\widehat{N}$ (称为 "极大似然估计") 由下述关系式所定义:

$$\widehat{N} = \left[\frac{Mn}{m}\right],$$

其中 $[\cdot]$ 表示整数部分 (本题的延伸见第 7 节第 4 题).

22. 在 (初等) 组合论中, 二项系数 $C_M^n \equiv \frac{(M)_n}{n!} = \frac{M!}{n!(M-n)!}$(也记为 $\binom{M}{n}$), 以及排列数 $(M)_n \equiv M(M-1)\cdots(M-n+1)$ 通常仅仅对正整数 $n, M \in \mathbb{N} = \{1,2,\cdots\}$ 有定义. 在分析学中, 考察 M 可以取任意实数 X 的排列数 $(M)_n$ 和二项系数 C_M^n 往往是有益的. 其中假定 $n \in \{0, \pm1, \pm2, \cdots\}$, 并且定义 $0!=1, (X)_0=1, C_X^0=1$,

$$(X)_n = X(X-1)\cdots(X-n+1), \quad C_X^n = \frac{(X)_n}{n!}, \quad 如果\ n>0,$$

而 $C_X^n=0$, 如果 $n<0$. 对所给出的这些定义 (再结合第 2 题中的一些关系式), 证明它们的下述性质 $(X, Y \in \mathbb{R},\ n \in \mathbb{Z} = \{0, \pm1, \pm2, \cdots\})$:

$$C_X^{n-1} + C_X^n = C_{X+1}^n \quad (帕斯卡三角形);$$

$$C_{X+Y}^n = \sum_{k=0}^n C_X^k C_Y^{n-k} \quad (范德蒙德二项卷积);$$

$$C_{X-1}^n = \sum_{k=0}^n (-1)^{n-k} C_X^k; \quad C_{n-X}^{n-m} = \sum_{k=0}^{n-m} (-1)^k C_X^k C_{n-k}^m;$$

$$C_{X+Y+n-1}^n = \sum_{k=0}^n C_{X+n-k-1}^{n-k} C_{Y+k-1}^k; \quad C_{-X}^n = (-1)^n C_{X+n-1}^n.$$

23. 罐中放有 N 个球, 其中有 n 个白球, $N-n$ 个黑球 $(n \geqslant 2)$. 从中进行无放回的取球. 我们来考察大小为 M 的样本. 以 $A_{i,j}$ 表示在第 i, j 个位置上放着白球的事件 $(1 \leqslant i < j \leqslant M)$, 以 $A_{i,j,k}$ 表示在第 i, j, k 个位置上放着白球的事件 $(1 \leqslant i < j < k \leqslant M)$. 试求事件 $A_{i,j}$ 与 $A_{i,j,k}$ 的概率.

24. 从一副 52 张扑克 (不含大小王牌) 中抽取 13 张牌, 以 P_n 表示其中恰有 n 张黑桃的概率. 试写出 P_n 的表达式.

25. 圆周上分布着 n 个点 $(n \geqslant 3)$, 从中随机选择两个点, 求此二点相邻的概率.

26. (夫妻邻座问题) n 对夫妻 $(n \geqslant 3)$ 围着一张圆桌坐下, 男士与女士均相间而坐. 试问, 有多少种坐法使得任何一对夫妻都不相邻?

提示: 将 $2n$ 个座位 (按顺时针方向) 依次编为 $1, 2, \cdots, 2n$ 号. 为确定起见, 假设在第 1 号位置上坐一位女士. 以 A_k 表示在第 k 号和第 $k+1$ 号位置上坐着某一对夫妻的事件 $(1 \leqslant k \leqslant 2n$, 将 $2n+1$ 号位置上理解为 1 号位置, 如此等等). 那么, 所要求的夫妻都不相邻的概率就是 $\mathbf{P}(\bigcap_{k=1}^{2n} \overline{A}_k)$, 按照容斥公式 (第 1 节第 12(b) 题) 有

$$\mathbf{P}\left(\bigcap_{k=1}^{2n} \overline{A}_k\right) = 1 - \mathbf{P}\left(\bigcup_{k=1}^{2n} A_k\right) = 1 - \sum_i \mathbf{P}(A_i) + \sum_{i<j} \mathbf{P}(A_i \cap A_j) - \cdots.$$

经计算可得, 对于 $1 \leqslant i \leqslant 2n$,

$$\mathbf{P}(A_i) = n\left(\frac{(n-1)!}{n!}\right)^2,$$

对于 $1 \leqslant i < j \leqslant 2n$,

$$\mathbf{P}(A_i \cap A_j) = \begin{cases} n(n-1)\left(\dfrac{(n-2)!}{n!}\right)^2, & 如果\ |i-j| \neq 1, \\ 0, & 如果\ |i-j| = 1, \end{cases}$$

其中 $\mathbf{P}(A_1 \cap A_{2n}) = 0$, 并且一般地, 对于 $i_1 < \cdots < i_k$,

$$\mathbf{P}(A_{i_1} \cap \cdots \cap A_{i_k}) = \begin{cases} \dfrac{n!}{(n-k)!}\left(\dfrac{(n-k)!}{n!}\right)^2, & \text{若 } 1 \leqslant j \leqslant k-1 \text{有} |i_{j+1} - i_j| \geqslant 2, \\ & \text{或 } 2n + i_1 - i_k \geqslant 2, \\ 0, & \text{其他情况}. \end{cases}$$

事实上, 所求的概率等于

$$\mathbf{P}\left(\bigcap_{k=1}^{2n} \overline{A}_k\right) = \sum_{k=0}^{n} (-1)^k \frac{(n-k)!}{n!} d_n^k,$$

其中 d_n^k 是可以选出 k 对互不相交的不同位置的选法数目, 互不相交的意思是: 如果 (i_1, i_2) 和 (i_3, i_4) 是其中的两对被选出的位置, 那么就有 $(i_1, i_2) \cap (i_3, i_4) = \varnothing$.

证明

$$d_n^k = C_{2n-k}^k \frac{2n}{2n-k}.$$

于是我们得到这样的答案: 任何一对夫妻都不相邻而坐的概率等于

$$\frac{1}{n!} \sum_{k=0}^{n} (-1)^k (n-k)! \frac{2n}{2n-k} C_{2n-k}^k.$$

27. (拉丁方) 我们来考察正整数 $1, 2, \cdots, n$ 以及由它们所列成的具有下述性质的 $n \times n$ 矩阵 (正方形): 每个数在每一行与每一列中都只出现一次. $n = 2$ 时的例子如下:

$$\begin{array}{|cc|} \hline 1 & 2 \\ 2 & 1 \\ \hline \end{array}, \qquad \begin{array}{|cc|} \hline 2 & 1 \\ 1 & 2 \\ \hline \end{array}.$$

$n = 3$ 时的例子则有:

$$\begin{array}{|ccc|} \hline 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \\ \hline \end{array}, \qquad \begin{array}{|ccc|} \hline 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \\ \hline \end{array}.$$

以 L_n 表示不同的 $n \times n$ 的拉丁方的数目, 证明

$$L_n \geqslant n!(n-1)! \cdots 1! \quad \left(= \prod_{k=1}^{n} k!\right).$$

注: 难以得到 L_n 的确切表达式 (个别的情况有: $L_2 = 2, L_3 = 12, L_4 = 576$). 现在知道其渐近表达式为

$$\ln L_n = n^2 \ln n + O(n^2), \quad n \to \infty.$$

28. (波利亚模型) 罐子里放有 r 个红球和 b 个黑球. 从中 "随机地" 取出一个球, 然后把它以及一个新的与它同色的球一起放回罐中. 再重复同样的 (取球和放球) 过程, 并一直进行下去.

以 S_n 表示在 n 次这样的取球中所取出的红球数目. 证明,

$$\mathbf{P}\{S_n = x\} = \frac{C_{r+x-1}^{r-1} C_{b+n-x-1}^{b-1}}{C_{r+b+n-1}^{n}}, \quad 0 \leqslant x \leqslant n.$$

29. 在上面的波利亚罐子模型问题中, 记

$$p = \frac{r}{r+b}, \quad q = \frac{b}{r+b} \quad \text{以及} \quad \gamma = \frac{1}{r+b}.$$

假设 $n \to \infty$, $p \to 0$, $\gamma \to 0$, 使得 $np \to \lambda$, $n\gamma \to 1/\rho$. 证明, 对于固定的 x, 在 $n \to \infty$ 时有

$$\mathbf{P}\{S_n = x\} \to C_{\lambda\rho+x-1}^{x} \left(\frac{\rho}{1+\rho}\right)^{\lambda\rho} \left(\frac{1}{1+\rho}\right)^{x}.$$

30. $2n$ 个球中, 黑球与白球各占一半, 我们来考察将它们放入 m 个盒子的问题. m 个盒子依次编有号码 $1, \cdots, m$. 每个黑球落入第 j 号盒子的概率为 p_j $(p_1 + \cdots + p_m = 1)$, 每个白球落入第 j 号盒子的概率为 q_j $(q_1 + \cdots + q_m = 1)$. 以 ν 记恰好落有 1 个黑球和 1 个白球的盒子的数目. 试求概率 $\mathbf{P}\{\nu = k\}$, $k = 0, 1, \cdots, m$ 和数学期望 $\mathbf{E}\nu$.

31. (关于斯特林公式, 还可参阅第 3 节第 16 题和第八章第 8 节第 1 题) 从数学分析中已经知道如下的斯特林准确公式

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \left(1 + \frac{1}{12n} + \frac{1}{288n^2} - \frac{139}{5140n^3} + O\left(\frac{1}{n^4}\right)\right).$$

试运用关系式

$$\ln n! = \sum_{k=2}^{n} \ln k \quad \text{和} \quad \ln(n-1)! < \int_1^n \ln t \, dt < \ln n!,$$

其中 $\int_1^n \ln t \, dt = n \ln n - n + 1$, 给出关于 $n!$ 的粗略的上下界估计:

$$e\left(\frac{n}{e}\right)^n < n! < en\left(\frac{n}{e}\right)^n, \tag{*}$$

并由其推出 (通常的) 斯特林公式

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

32. (关于调和数的渐近展开) 所谓调和数, 是指 $H_n = \sum_{k=1}^{n} \frac{1}{k}$, $n \geqslant 1$.

由数学分析中知道

$$H_n = \ln n + \gamma + \frac{1}{2n} - \frac{1}{12n^2} + \frac{1}{120n^4} + O\left(\frac{1}{n^6}\right),$$

其中 $\gamma = 0.5772\cdots$, 称为欧拉常数.

试运用前一题中的通过积分估计和数的办法, 证明, 对于一切 $n \geqslant 1$, 都有

$$\ln n + \frac{1}{n} \leqslant H_n \leqslant \ln n + 1$$

(并且由此有 $\lim\limits_{n\to\infty}(\frac{H_n}{\ln n}) = 1$).

§3. 条件概率. 独立性

1. 试举例说明, 下列等式一般来说是不成立的:

$$\mathbf{P}(B|A) + \mathbf{P}(B|\overline{A}) = 1,$$
$$\mathbf{P}(B|A) + \mathbf{P}(\overline{B}|\overline{A}) = 1.$$

2. 罐中放有 M 个球, 其中 M_1 个为白球. 考察从中取出的大小为 n 的抽样. 以 B_j 表示在第 j 步所取出的球为白球的事件, 以 A_k 表示在所取出的一共 n 个球中恰有 k 个白球的事件. 证明, 无论是有放回的抽样, 还是无放回的抽样, 都有

$$\mathbf{P}(B_j|A_k) = \frac{k}{n}.$$

提示: 对于有放回的抽样, 应当首先证明

$$\mathbf{P}(B_j \cap A_k) = \frac{C_{n-1}^{k-1} M_1^k (M-M_1)^{n-k}}{M^n},$$
$$\mathbf{P}(A_k) = \frac{C_n^k M_1^k (M-M_1)^{n-k}}{M^n}.$$

而对于无放回的抽样, 则应当首先证明

$$\mathbf{P}(B_j \cap A_k) = \frac{C_{n-1}^{k-1} (M_1)_k (M-M_1)_{n-k}}{(M)_n},$$
$$\mathbf{P}(A_k) = \frac{C_n^k (M_1)_k (M-M_1)_{n-k}}{(M)_n}.$$

3. 设 $A_1, \cdots, A_n$ 为相互独立的事件, 其中 $\mathbf{P}(A_i) = p_i$.

(a) 证明

$$\mathbf{P}\left(\bigcup_i^n A_i\right) = 1 - \prod_{i=1}^n \mathbf{P}(\overline{A}_i). \qquad (*)$$

(b) 设 P_0 为事件 $A_1, \cdots, A_n$ 中任何一个都不发生的概率. 证明,

$$P_0 = \prod_{i=1}^n (1-p_i).$$

提示: 如果首先证明若事件 $A_1, \cdots, A_n$ 相互独立, 则事件 $\widetilde{A}_1, \cdots, \widetilde{A}_n$ 亦相互独立, 其中 $\widetilde{A}_i$ 表示 A_i 或 $\overline{A}_i$, 那么就可以给出一个关于 $(*)$ 的直接证明 (无需转化为第 1 节第 12 题中的容斥公式).

4. 设 A 和 B 为相互独立的事件, 试用 $\mathbf{P}(A)$ 和 $\mathbf{P}(B)$ 表达 A 和 B 之中恰有 k 个, 至少有 k 个, 最多有 k 个发生的概率, 其中 $k = 0, 1, 2$. 并与第 1 节第 13 题相比较.

5. 设事件 A 与自己独立, 即 A 与 A 相互独立. 证明 $\mathbf{P}(A)$ 等于 0 或 1.

证明, 如果事件 A 与 B 相互独立, 且 $A \subseteq B$, 则或者 $\mathbf{P}(A) = 0$, 或者 $\mathbf{P}(B) = 1$.

6. 证明, 如果 A 为事件, 且 $\mathbf{P}(A)$ 等于 0 或 1, 则 A 与任何事件 B 独立.

7. 观察图 4 (《概率》第一卷第一章第 31 页的附图) 中的电路图, 假设其中的元件 A, B, C, D 和 E 中的每一个处于断开状态 (即不工作状态, 不能通过信号) 和处于闭合状态 (即工作状态, 允许信号通过) 的概率都分别是 p 和 q, 并且它们之间的工作状态相互独立. 试问, 如果在入口处输入一个信号, 那么能够在出口处接收到该信号的概率为多少? 如果能够在出口处接收到信号, 那么元件 E 处于断开状态的条件概率是多少?

提示: 以 S 表示可以在出口处接收到由入口处输入的信号的事件. 证明

$$\mathbf{P}(S|E) = 1 - 2p^2 + p^4,$$
$$\mathbf{P}(S|\overline{E}) = 2q^2 - q^4;$$

并利用全概率公式

$$\mathbf{P}(S) = q(1-p^2)^2 + pq^2(2-q^2).$$

再验证

$$\mathbf{P}(E|S) = \frac{(1-p^2)^2}{(1-p^2)^2 + pq(2-q^2)}.$$

8. 设 $\mathbf{P}(A+B) > 0$, 证明

$$\mathbf{P}(A|A+B) = \frac{\mathbf{P}(A)}{\mathbf{P}(A+B)}.$$

9. 设事件 A 与事件 $B_n\,(n \geqslant 1)$ 独立, 且有 $B_i \cap B_j = \varnothing, i \neq j$. 证明, A 与事件 $\bigcup\limits_{n=1}^{\infty} B_n$ 独立.

10. 证明, 如果 $\mathbf{P}(A|C) > \mathbf{P}(B|C)$, $\mathbf{P}(A|\overline{C}) > \mathbf{P}(B|\overline{C})$, 则 $\mathbf{P}(A) > \mathbf{P}(B)$.

11. 证明

$$\mathbf{P}(A|B) = \mathbf{P}(A|BC)\mathbf{P}(C|B) + \mathbf{P}(A|B\overline{C})\mathbf{P}(\overline{C}|B).$$

12. 设 X 与 Y 是相互独立的参数为 n 和 p 的二项分布随机变量, 证明

$$\mathbf{P}(X = k|X+Y=m) = \frac{C_n^k C_n^{m-k}}{C_{2n}^m}, \quad k = 0, 1, \cdots, \min(m, n).$$

13. 设 A, B, C 为两两独立的等概事件, 且 $A \cap B \cap C = \varnothing$. 试求 $\mathbf{P}(A)$ 的最大可能值.

14. 罐中装着 1 个白球, 往里面放入一个 "胡乱选出的" 球, 可能是白球, 也可能是黑球 (以相同的概率选择). 此后以随机的方式从罐中取出一个球, 发现该球为白球. 试求此时罐中剩下的球也是白球的条件概率.

15. 如果事件 A 与事件 B 相互独立, 则按照定义就有 $\mathbf{P}(AB) = \mathbf{P}(A)\mathbf{P}(B)$. 试描述对任意事件 A 与 B, 如下两个不等式成立的条件:

$$\mathbf{P}(AB) \leqslant \mathbf{P}(A)\mathbf{P}(B) \quad 与 \quad \mathbf{P}(AB) \geqslant \mathbf{P}(A)\mathbf{P}(B).$$

16. 试利用广义斯特林公式 ($n! \sim \sqrt{2\pi n}n^n e^{-n}$, $n \to \infty$), 证明, Γ 函数 $\Gamma(\nu) = \int_0^\infty u^{\nu-1}e^{-u}du$, $\nu > 0$, 具有性质

$$\Gamma(\nu) \sim \sqrt{2\pi\nu}\nu^\nu e^{-\nu}, \quad \nu \to \infty.$$

§4. 随机变量及其特征

在本章中, 空间 Ω 由有穷个元素构成, 因此, 此处所考察的所有随机变量也都只取有穷个可能值.

1. 验证示性函数 $I_A = I_A(\omega)$ 的下列性质:

$$I_\varnothing = 0, \quad I_\Omega = 1, \quad I_{\overline{A}} = 1 - I_A,$$
$$I_{AB} = I_A \cdot I_B, \quad I_{A\cup B} = I_A + I_B - I_{AB},$$
$$I_{A\setminus B} = I_A(1 - I_B), \quad I_{A\triangle B} = (I_A - I_B)^2 = I_A + I_B \pmod 2,$$
$$I_{\bigcup\limits_{i=1}^n A_i} = 1 - \prod_{i=1}^n (1 - I_{A_i}), \quad I_{\overline{\bigcup\limits_{i=1}^n A_i}} = \prod_{i=1}^n (1 - I_{A_i}), \quad I_{\sum\limits_{i=1}^n A_i} = \sum_{i=1}^n I_{A_i},$$

其中 $A \triangle B$ 是集合 A 与 B 的对称差, 亦即集合 $(A\setminus B) \cup (B\setminus A)$, 和 $\sum$ 表示不相交事件的并 ($\cup$).

2. 利用第 1 题中的命题, 证明关于三个事件 A, B, C 的示性函数的容斥公式:

$$I_{A\cup B\cup C} = I_A + I_B + I_C - \left(I_{A\cap B} + I_{A\cap C} + I_{B\cap C}\right) + I_{A\cap B\cap C}.$$

写出关于 n 个事件 $A_1, \cdots, A_n$ 之并的示性函数 $I_{A_1\cup\cdots\cup A_n}$ 的相应形式的表达式.

3. 设 $\xi_1, \cdots, \xi_n$ 为相互独立的伯努利随机变量, 有

$$\mathbf{P}\{\xi_i = 0\} = 1 - \lambda_i\Delta,$$
$$\mathbf{P}\{\xi_i = 1\} = \lambda_i\Delta,$$

其中 Δ 为一个很小的数, $\Delta > 0$, $\lambda_i > 0$. 证明

$$\mathbf{P}\{\xi_1 + \cdots + \xi_n = 1\} = \left(\sum_{i=1}^n \lambda_i\right)\Delta + O(\Delta^2),$$
$$\mathbf{P}\{\xi_1 + \cdots + \xi_n > 1\} = O(\Delta^2).$$

4. 证明 $\inf\limits_{-\infty<a<\infty} \mathbf{E}(\xi - a)^2$ 在 $a = \mathbf{E}\xi$ 时达到, 由此有

$$\inf_{-\infty<a<\infty} \mathbf{E}(\xi - a)^2 = \mathbf{D}\xi.$$

提示: 设 $\mathbf{E}\xi = 0$, 可得 $\mathbf{E}(\xi - a)^2 = \mathbf{D}\xi + a^2 \geqslant \mathbf{D}\xi$.

5. 设 ξ 为随机变量, 具有分布函数 $F_\xi(x) = \mathbf{P}(\xi \leqslant x)$. 称 $\mu = \mu(\xi)$(或 $\mu = \mu(F_\xi)$) 为随机变量 ξ (或分布函数 $F_\xi(x)$) 的中位数, 如果

$$F_\xi(\mu-) \leqslant \frac{1}{2} \leqslant F_\xi(\mu).$$

(关于中位数的另一定义见下面的第 23 题.)

证明

$$\inf_{-\infty<a<\infty} \mathbf{E}|\xi - a| = \mathbf{E}|\xi - \mu|.$$

提示: 设 $\mu = 0$, 先考虑 $a > 0$ 的情形, 此时我们有

$$\mathbf{E}|\xi - a| = \mathbf{E}|\xi| + \mathbf{E}f(\xi),$$

其中

$$f(x) = \begin{cases} a, & x \leqslant 0, \\ a - 2x, & 0 < x < a, \\ -a, & x \geqslant a. \end{cases}$$

既然 $f(x) \geqslant 0$, 所以 $\mathbf{E}f(\xi) \geqslant 0$[①], $\mathbf{E}|\xi - a| \geqslant \mathbf{E}|\xi|$. 可类似证明 $a < 0$ 的情形.

6. 设 $P_\xi(x) = \mathbf{P}\{\xi = x\}$, $F_\xi(x) = \mathbf{P}\{\xi \leqslant x\}$. 证明, 对于 $a > 0$, $-\infty < b < \infty$, 有

$$P_{a\xi+b}(x) = P_\xi\left(\frac{x-b}{a}\right),$$
$$F_{a\xi+b}(x) = F_\xi\left(\frac{x-b}{a}\right).$$

并证明, 对于 $y \geqslant 0$, 有

$$F_{\xi^2}(y) = F_\xi(+\sqrt{y}) - F_\xi(-\sqrt{y}) + P_\xi(-\sqrt{y});$$

而对于 $\xi^+ = \max(\xi, 0)$, 有

$$F_{\xi^+}(x) = \begin{cases} 0, & x < 0, \\ F_\xi(x), & x \geqslant 0. \end{cases}$$

提示: 我们注意:

$$\{a\xi + b = x\} = \left\{\xi = \frac{x-b}{a}\right\}, \quad \{a\xi + b \leqslant x\} = \left\{\xi \leqslant \frac{x-b}{a}\right\},$$
$$\{\xi^2 \leqslant y\} = \{\xi = -\sqrt{y}\} \cup \{\xi \leqslant +\sqrt{y}\} \backslash \{\xi \leqslant -\sqrt{y}\},$$
$$\{\xi^+ \leqslant x\} = \begin{cases} \varnothing, & x < 0, \\ \{\xi \leqslant x\}, & x \geqslant 0. \end{cases}$$

7. 设 ξ 与 η 为两个随机变量, $\mathbf{D}\xi > 0$, $\mathbf{D}\eta > 0$, 以 $\rho = \rho(\xi, \eta)$ 表示它们的相关系数. 证明 $|\rho| \leqslant 1$, 并且 $|\rho| = 1$ 当且仅当可以找到常数 a 与 b, 使得 $\eta = a\xi + b$. 更进一步, 如果 $\rho = 1$, 则有

$$\frac{\eta - \mathbf{E}\eta}{\sqrt{\mathbf{D}\eta}} = \frac{\xi - \mathbf{E}\xi}{\sqrt{\mathbf{D}\xi}}$$

(这表明当 $\rho = 1$ 时, 常数 $a > 0$), 如果 $\rho = -1$, 则有

$$\frac{\eta - \mathbf{E}\eta}{\sqrt{\mathbf{D}\eta}} = -\frac{\xi - \mathbf{E}\xi}{\sqrt{\mathbf{D}\xi}}$$

(这表明当 $\rho = -1$ 时, 常数 $a < 0$).

[①] 由于 $\mu = 0$, $a > 0$, 所以 $\mathbf{P}\{\xi \leqslant 0\} \geqslant \frac{1}{2}$, $\mathbf{P}\left\{\xi > \frac{a}{2}\right\} < \frac{1}{2}$, 从而有

$$\begin{aligned}
\mathbf{E}f(\xi) &= aP(\xi \leqslant 0) + \mathbf{E}f(\xi)I(0 < \xi < a) - aP(\xi \geqslant a) \\
&\geqslant aP(\xi \leqslant 0) + \mathbf{E}(a - 2\xi)I(0 < \xi \leqslant \frac{a}{2}) + \mathbf{E}(a - 2\xi)I(\frac{a}{2} < \xi < a) - aP(\xi \geqslant a) \\
&\geqslant aP(\xi \leqslant 0) - aP\left(\xi > \frac{a}{2}\right) + \mathbf{E}(a - 2\xi)I(0 < \xi \leqslant \frac{a}{2}) \\
&= a\left(P(\xi \leqslant 0) - P\left(\xi > \frac{a}{2}\right)\right) + \mathbf{E}(a - 2\xi)I(0 < \xi \leqslant \frac{a}{2}) \\
&\geqslant \mathbf{E}(a - 2\xi)I(0 < \xi \leqslant \frac{a}{2}) \geqslant 0
\end{aligned}$$

—— 译者注.

8. 设 ξ 与 η 为两个随机变量, $\mathbf{E}\xi = \mathbf{E}\eta = 0$, $\mathbf{D}\xi = \mathbf{D}\eta = 1$, 以 $\rho = \rho(\xi, \eta)$ 表示它们的相关系数. 证明,

$$\mathbf{E}\max(\xi^2, \ \eta^2) \leqslant 1 + \sqrt{1 - \rho^2}.$$

提示: 利用等式

$$\max(\xi^2, \ \eta^2) = \frac{1}{2}(\xi^2 + \eta^2 + |\xi^2 - \eta^2|)$$

和柯西–布尼亚科夫斯基不等式.

9. 利用第 1 题关于示性函数的性质中的 $I_{\bigcup_{i=1}^n A_i} = 1 - \prod_{i=1}^n (1 - I_{A_i})$, 证明关于事件之并的概率容斥公式:

$$\begin{aligned}
\mathbf{P}(A_1 \cup \cdots \cup A_n) = &\sum_{1 \leqslant i_1 \leqslant n} \mathbf{P}(A_{i_1}) - \sum_{1 \leqslant i_1 < i_2 \leqslant n} \mathbf{P}(A_{i_1} \cap A_{i_2}) + \cdots \\
&+ (-1)^{m+1} \sum_{1 \leqslant i_1 < \cdots < i_m \leqslant n} \mathbf{P}(A_{i_1} \cap \cdots \cap A_{i_m}) + \cdots + (-1)^{n+1}\mathbf{P}(A_1 \cap \cdots \cap A_n)
\end{aligned}$$

(与第 1 节第 12 题相比较).

提示: 记 $X_i = I_{A_i}$, 先证明如下的关于示性函数的容斥公式:

$$\begin{aligned}
1 - \prod_{i=1}^n (1 - X_i) = &\sum_{1 \leqslant i \leqslant n} X_i - \sum_{1 \leqslant i_1 < i_2 \leqslant n} X_{i_1}X_{i_2} + \cdots + \\
&+ (-1)^{m+1} \sum_{1 \leqslant i_1 < \cdots < i_m \leqslant n} X_{i_1} \cdots X_{i_m} + \cdots + (-1)^{n+1} X_1 \cdots X_n.
\end{aligned}$$

再利用关系式 $\mathbf{P}\left(\bigcup_{i=1}^n I_{A_i}\right) = \mathbf{E}I_{\bigcup_{i=1}^n A_i}$ (比较第 1 节第 13 题的提示).

10. 设 $\xi_1, \cdots, \xi_n$ 为相互独立的随机变量, $\varphi_1 = \varphi_1(\xi_1, \cdots, \xi_k)$ 和 $\varphi_2 = \varphi_2(\xi_{k+1}, \cdots, \xi_n)$ 分别是两个关于 $\xi_1, \cdots, \xi_k$ 和 $\xi_{k+1}, \cdots, \xi_n$ 的函数, 则随机变量 φ_1 与 φ_2 相互独立. 试证明之.

11. 证明, 随机变量 $\xi_1, \cdots, \xi_n$ 相互独立, 当且仅当对一切 $x_1, \cdots, x_n$, 都有

$$F_{\xi_1, \cdots, \xi_n}(x_1, \cdots, x_n) = F_{\xi_1}(x_1) \cdots F_{\xi_n}(x_n),$$

其中 $F_{\xi_1, \cdots, \xi_n}(x_1, \cdots, x_n) = \mathbf{P}\{\xi_1 \leqslant x_1, \cdots, \xi_n \leqslant x_n\}$.

12. 证明, 随机变量 ξ 与自己独立 (即 ξ 与 ξ 相互独立), 当且仅当 $\xi(\omega) \equiv \text{const}$, $\omega \in \Omega$.

13. 当随机变量 ξ 满足怎样的条件时, 有 ξ 与 $\sin\xi$ 相互独立?

14. 设 ξ 与 η 为相互独立的随机变量, 且 $\eta \neq 0$. 试利用概率 $\mathbf{P}\{\xi \leqslant x\}$ 与 $\mathbf{P}\{\eta \leqslant y\}$, $x, y \in \mathbb{R}$, 表示概率 $\mathbf{P}\{\xi\eta \leqslant z\}$ 与 $\mathbf{P}\left\{\frac{\xi}{\eta} \leqslant z\right\}$.

15. 设 ξ, η, ζ 为随机变量, 有 $|\xi| \leqslant 1, |\eta| \leqslant 1, |\zeta| \leqslant 1$. 证明如下的贝尔不等式: (参阅 [123])

$$|\mathbf{E}\,\xi\zeta - \mathbf{E}\,\eta\zeta| \leqslant 1 - \mathbf{E}\,\xi\eta.$$

提示: 利用 $\xi(1+\eta) \leqslant 1+\eta$.

16. 以相互独立的方式将 k 个球放入 n 个罐子 (每个球放入每个罐子的概率都是 $1/n$), 试求不空的罐子的数目的数学期望.

17. 设 $\xi_1, \cdots, \xi_n$ 为独立同分布的随机变量, 有 $\mathbf{P}\{\xi_1 = 1\} = p$, $\mathbf{P}\{\xi_1 = 0\} = 1-p$, 其中 $0 < p < 1$. 记 $S_k = \xi_1 + \cdots + \xi_k$, $k \leqslant n$. 证明, 对于 $1 \leqslant m \leqslant n$, 有

$$\mathbf{P}(S_m = k | S_n = l) = \frac{C_m^k C_{n-m}^{l-k}}{C_n^l}.$$

18. 设 $\xi_1, \cdots, \xi_n$ 为相互独立的随机变量, 而

$$\xi_{\min} = \min(\xi_1, \cdots, \xi_n), \quad \xi_{\max} = \max(\xi_1, \cdots, \xi_n).$$

证明

$$\mathbf{P}\{\xi_{\min} \geqslant x\} = \prod_{i=1}^{n} \mathbf{P}\{\xi_i \geqslant x\}, \quad \mathbf{P}\{\xi_{\max} < x\} = \prod_{i=1}^{n} \mathbf{P}\{\xi_i < x\}.$$

19. 设[①] $\xi_1, \cdots, \xi_{2n}$ 为独立同分布的随机变量, 有 $\mathbf{P}\{\xi_1 = 1\} = \mathbf{P}\{\xi_1 = -1\} = \frac{1}{2}$, 记 $S_{2n} = \xi_1 + \cdots + \xi_{2n}$, $M_{2n} = \max(S_1, \cdots, S_{2n})$. 证明, 对于 $k \leqslant n$

$$\mathbf{P}\{M_{2n} \geqslant k,\ S_{2n} = 0\} = \mathbf{P}\{S_{2n} = 2k\},$$

因而

$$\mathbf{P}(M_{2n} \geqslant k | S_{2n} = 0) = \frac{\mathbf{P}\{S_{2n} = 2k\}}{\mathbf{P}\{S_{2n} = 0\}} = \frac{C_{2n}^{n+k}}{C_{2n}^{n}}.$$

并由此推出

$$\mathbf{E}(M_{2n} | S_{2n} = 0) = \frac{1}{2}\left(\frac{1}{\mathbf{P}\{S_{2n} = 0\}} - 1\right).$$

20. 举例说明, 存在这样的两个随机变量 ξ 与 η, 它们具有相同的分布函数 $(F_\xi = F_\eta)$, 但却有 $\mathbf{P}\{\xi \neq \eta\} > 0$.

21. 设 ξ, η 与 ζ 均为随机变量, 并且 ξ 与 η 的分布相同. 试问, $\xi\zeta$ 与 $\eta\zeta$ 的分布是否一定相同?

22. 举例说明, 存在这样的两个随机变量 ξ 与 η, 它们自己不相互独立, 而 ξ^2 与 η^2 却相互独立.

① 原文为 设 $S_{2n} = \xi_1 + \cdots + \xi_{2n}$, $M_{2n} = \max(S_1, \cdots, S_{2n})$. 证明, $\cdots\cdots$. 此处关于随机变量的独立性和分布的假设系译者顺应题意所加 —— 译者注.

23. 设 ξ 为离散随机变量, 考察如下三种定义随机变量 ξ 的中位数 $\mu = \mu(\xi)$ 的方法 (参阅第 5 题):

(a) $\max\Big(\mathbf{P}\{\xi > \mu\}, \mathbf{P}\{\xi < \mu\}\Big) \leqslant \frac{1}{2}$,

(b) $\mathbf{P}\{\xi < \mu\} \leqslant \frac{1}{2} \leqslant \mathbf{P}\{\xi \leqslant \mu\}$,

(c) $\mu = \inf\Big\{x \in \mathbb{R}:\ \mathbf{P}\{\xi \leqslant x\} \geqslant \frac{1}{2}\Big\}$.

假设 M_a, M_b, M_c 是分别按照定义 (a),(b) 和 (c) 所确定的中位数集合. 试问, 它们之间存在何种关系?

24. 罐中放有 N 个球, 其中有 a 个白球, b 个黑球和 c 个红球, $a+b+c=N$. 从中取出 n 个球, 设它们中有 ξ 个白球和 η 个黑球. 证明, 如果所进行的取球是 "有放回抽球", 则

$$\mathbf{cov}(\xi, \eta) = -npq,$$

其中 $p = a/N$, $q = b/N$. 如果所进行的取球是 "无放回抽球", 则

$$\mathbf{cov}(\xi, \eta) = -npq\frac{N-n}{N-1}.$$

证明, 在两种情况下都有

$$\rho(\xi, \eta) = -\sqrt{\frac{pq}{(1-p)(1-q)}}.$$

§5. 伯努利概型 I. 大数定律

1. 设 ξ 与 η 是两个具有相关系数 ρ 的随机变量. 证明如下的二维切比雪夫不等式:

$$\mathbf{P}\Big\{|\xi - \mathbf{E}\xi| \geqslant \varepsilon\sqrt{\mathbf{D}\xi} \text{ 或 } |\eta - \mathbf{E}\eta| \geqslant \varepsilon\sqrt{\mathbf{D}\eta}\Big\} \leqslant \frac{1}{\varepsilon^2}(1 + \sqrt{1-\rho^2}).$$

提示: 不失一般性, 可设 $\mathbf{E}\xi = \mathbf{E}\eta = 0$, $\mathbf{D}\xi = \mathbf{D}\eta = 1$, 于是

$$\mathbf{P}\Big\{|\xi| \geqslant \varepsilon \text{ 或 } |\eta| \geqslant \varepsilon\Big\} = \mathbf{P}\Big\{\max(\xi^2, \eta^2) \geqslant \varepsilon^2\Big\}.$$

再利用 (通常的) 切比雪夫不等式和第 4 节第 8 题中的不等式即可.

2. 设 $f(x)$ 为非负偶函数, 在正半轴上非降. 证明, 如果对于随机变量 $\xi = \xi(\omega)$ 有 $\xi(\omega) \leqslant C$, 其中 $C > 0$, 则有如下的下界估计:

$$\mathbf{P}\{|\xi| \geqslant \varepsilon\} \geqslant \frac{\mathbf{E}f(\xi) - f(\varepsilon)}{f(C)}.$$

特别地, 对于 $f(x) = x^2$, 有

$$\frac{\mathbf{E}\xi^2 - \varepsilon^2}{C^2} \leqslant \mathbf{P}\{|\xi - \mathbf{E}\xi| \geqslant \varepsilon\} \left(\leqslant \frac{\mathbf{D}\xi}{\varepsilon^2}\right).$$

提示: 利用下述关系式

$$\mathbf{E}f(\xi) \leqslant \mathbf{E}f(|\xi|) \leqslant f(C)\mathbf{P}\{|\xi| \geqslant \varepsilon\} + f(\varepsilon).$$

3. 设 $\xi_1, \cdots, \xi_n$ 为相互独立的随机变量, 有 $\mathbf{D}\xi_i \leqslant C$. 证明

$$\mathbf{P}\left\{\left|\frac{\xi_1 + \cdots + \xi_n}{n} - \frac{\mathbf{E}(\xi_1 + \cdots + \xi_n)}{n}\right| \geqslant \varepsilon\right\} \leqslant \frac{C}{n\varepsilon^2}.$$

(该不等式表明, 大数定律可在更为广泛的情况下成立, 而不仅仅局限于伯努利情形.)

4. 设 $\xi_1, \cdots, \xi_n$ 为相互独立的伯努利随机变量, 有 $\mathbf{P}\{\xi_i = 1\} = p > 0$, $\mathbf{P}\{\xi_i = -1\} = 1 - p$. 证明有如下的伯恩斯坦估计成立: 存在 $a > 0$, 使得

$$\mathbf{P}\left\{\left|\frac{S_n}{n} - (2p-1)\right| \geqslant \varepsilon\right\} \leqslant 2e^{-a\varepsilon^2 n},$$

其中 $S_n = \xi_1 + \cdots + \xi_n$, 而 $\varepsilon > 0$.

提示: 参阅《概率》第一卷第 6 节中关于 (42) 式的证明.

5. 设 ξ 为非负随机变量, $a > 0$. 试分别在下列三种情形下求概率 $\mathbf{P}\{\xi \geqslant a\}$ 的最大可能值:

(i) 已知 $\mathbf{E}\,\xi = m$.

(ii) 已知 $\mathbf{E}\,\xi = m$, $\mathbf{D}\,\xi = \sigma^2$.

(iii) 已知 $\mathbf{E}\,\xi = m$, $\mathbf{D}\,\xi = \sigma^2$ 且 ξ 关于自己的均值 m 对称.

6. 设 $\xi_1, \cdots, \xi_N$ 为相互独立的伯努利随机变量, 有 $\mathbf{P}\{\xi_n = 1\} = p > 0$, $\mathbf{P}\{\xi_n = 0\} = q$, $n \leqslant N$. 令 $S_0 = 0$, $S_n = \xi_1 + \cdots + \xi_n$, $n \leqslant N$, 记 $P_n(k) = \mathbf{P}\{S_n = k\}$. 证明, 对于 $n < N$ 和 $k \geqslant 1$, 有

$$P_{n+1}(k) = pP_n(k-1) + qP_n(k).$$

7. 设 $\xi_1, \cdots, \xi_N$ 为相互独立的伯努利随机变量, 有 $\mathbf{P}\{\xi_i = 1\} = \mathbf{P}\{\xi_i = -1\} = 1/2$, $i = 1, \cdots, N$. 令 $S_m = \xi_1 + \cdots + \xi_m$. 证明, 对于 $2m \leqslant N$, 有

$$\mathbf{P}\{S_1 \cdots S_{2m} \neq 0\} = 2^{-2m} C_{2m}^m.$$

8. 设 M 个盒子被依次编为 $1, \cdots, M$ 号. 在第 n 号盒子中放有 1 个白球和 n 个黑球. 从每个盒中 "随机地" 取出一个球. 令

$$\xi_n = \begin{cases} 1, & \text{如果自第 } n \text{ 号盒子中取出的是白球}, \\ 0, & \text{如果自第 } n \text{ 号盒子中取出的是黑球}, \end{cases}$$

并令 $S_M = \xi_1 + \cdots + \xi_M$ 为所取出的白球总数. 证明, "S_M 在 M 很大时具有 $\ln M$ 的阶", 意即对任何 $\varepsilon > 0$, 当 $M \to \infty$ 时, 都有

$$\mathbf{P}\left\{\left|\frac{S_M}{\ln M} - 1\right| \geqslant \varepsilon\right\} \to 0$$

(采用推导 (8) 式时的术语).

9. 设 $\xi_1, \cdots, \xi_n$ 为相互独立的伯努利随机变量, 有 $\mathbf{P}\{\xi_k = 1\} = p_k$, $\mathbf{P}\{\xi_k = 0\} = 1 - p_k$, $1 \leqslant k \leqslant n$, 其中 $\frac{1}{n}\sum\limits_{k=1}^{n} p_k = a$. 记 $S_n = \xi_1 + \cdots + \xi_n$, 证明, 对于给定的 $0 < a < 1$, 方差 $\mathbf{D}S_n$ 在 $p_1 = \cdots = p_n = a$ 时达到其最大值.

10. 设 $\xi_1, \cdots, \xi_n$ 为相互独立的伯努利随机变量, 有 $\mathbf{P}\{\xi_k = 1\} = p$, $\mathbf{P}\{\xi_k = 0\} = 1 - p$, $1 \leqslant k \leqslant n$. 试在 n 次试验中仅成功了 1 次的条件下求第一次成功发生在第 m 次试验中的条件概率.

11. 设 $(p_1, \cdots, p_r)$ 与 $(q_1, \cdots, q_r)$ 为两个概率分布. 试证明吉布斯不等式:

$$-\sum_{i=1}^{r} p_i \ln p_i \leqslant -\sum_{i=1}^{r} p_i \ln q_i.$$

(由此可以推出熵 $H = -\sum\limits_{i=1}^{r} p_i \ln p_i \leqslant \ln r$, 参阅第 4 节中的有关说明.)

12. 在第 10 题的条件下, 证明雷尼 (Rényi) 不等式:

$$\mathbf{P}\left\{\left|\frac{S_n}{n} - p\right| \geqslant \varepsilon\right\} \leqslant \exp\left\{-\frac{n\varepsilon^2}{2pq(1 + \varepsilon/(2pq)^2)}\right\}.$$

§6. 伯努利概型 II. 极限定理 (棣莫弗–拉普拉斯局部定理、泊松定理)

1. 设 $n = 100$, $p = 1/10$, $2/10$, $3/10$, $4/10$, $5/10$. 试利用二项分布和泊松分布的统计表 (参阅 [12]), 将下列概率值与正态逼近和泊松逼近所给出的值进行比较:

$$\mathbf{P}\{10 < S_{100} \leqslant 12\}, \qquad \mathbf{P}\{20 < S_{100} \leqslant 22\},$$
$$\mathbf{P}\{33 < S_{100} \leqslant 35\}, \qquad \mathbf{P}\{40 < S_{100} \leqslant 42\},$$
$$\mathbf{P}\{50 < S_{100} \leqslant 52\}.$$

(可以利用计算机等工具.)

2. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布的伯努利随机变量序列, 有 $\mathbf{P}\{\xi_1 = 1\} = \mathbf{P}\{\xi_1 = 0\} = \frac{1}{2}$, $S_n = \xi_1 + \cdots + \xi_n$, $n \geqslant 1$. 令 $Z_n = 2S_n - n$ 表示 n 重伯努利试验中成功次数与失败次数之差. 证明[①]

$$\sup_j \left|\sqrt{\pi n}\,\mathbf{P}\{Z_{2n} = j\} - e^{-j^2/4n}\right| \to 0, \quad n \to \infty.$$

① 原书为: $\sup\limits_j \left|\sqrt{\pi n}\,\mathbf{P}\{Z_{2n} = j\} - e^{-j^2/4^n}\right| \to 0, \quad n \to \infty$ —— 译者注.

提示: 令 $2n=m$, $k=\frac{j}{2}+n$, 易见, 所要证明的命题可归结为证明 (其中 $p=1/2$):

$$\sup_k\left|\sqrt{\frac{\pi m}{2}}\mathbf{P}\{S_m=k\}-e^{-\frac{(k-mp)^2}{2mpq}}\right|\left(\equiv\sup_k\varepsilon(k,m)\right)\to 0,\quad m\to\infty.$$

再利用

$$\sup_k\varepsilon(k,m)=\max(a_m,\ b_m),$$

其中

$$a_m=\sup_{\{k:|k-mp|\leqslant(mpq)^s\}}\varepsilon(k,m),\quad b_m=\sup_{\{k:|k-mp|>(mpq)^s\}}\varepsilon(k,m),$$

此处 s 为区间 $(1/2,\ 3/2)$ 中的某个常数. 再证明当 $m\to\infty$ 时, 有 $a_m\to 0$, $b_m\to 0$.

3. 证明, 在泊松定理中, 有如下估计式成立 ($p=\lambda/n$):

$$\sup_k\left|P_n(k)-\frac{\lambda^k e^{-\lambda}}{k!}\right|\leqslant\frac{\lambda^2}{n}.$$

提示: 设 $\eta_1,\cdots,\eta_n$ 为独立同分布的参数为 λ/n 的泊松随机变量, 而 $\zeta_1,\cdots,\zeta_n$ 为独立同分布的伯努利随机变量, 有

$$\mathbf{P}\{\zeta_i=0\}=e^{\lambda/n}\left(1-\frac{\lambda}{n}\right),\quad \mathbf{P}\{\zeta_i=1\}=1-e^{\lambda/n}\left(1-\frac{\lambda}{n}\right).$$

令

$$\xi_i=\begin{cases}0, & \text{如果 } \eta_i=0,\ \zeta_i=0,\\ 1, & \text{其他情形,}\end{cases}$$

则易见 $\xi_1,\cdots,\xi_n$ 为独立同分布的伯努利随机变量, 有

$$\mathbf{P}\{\xi_1=0\}=1-\frac{\lambda}{n},\quad \mathbf{P}\{\xi_1=1\}=\frac{\lambda}{n},$$

从而 $\xi=\xi_1+\cdots+\xi_n$ 的分布为 $\mathbf{P}\{\xi=k\}=P_n(k)$. 再注意到随机变量 $\eta=\eta_1+\cdots+\eta_n$ 服从参数为 λ 的泊松分布, 并且对每个 $k=0,1,2,\cdots$, 都有

$$\left|\mathbf{P}\{\xi=k\}-\mathbf{P}\{\eta=k\}\right|\leqslant\mathbf{P}\{\xi\neq\eta\}\leqslant\frac{\lambda^2}{n}.$$

(试与第三章第 12 节中的证明与结论进行比较.)

4. 设 $\xi_1,\cdots,\xi_n$ 为独立同分布的随机变量, 有 $\mathbf{P}\{\xi_k=1\}=\mathbf{P}\{\xi_k=-1\}=\frac{1}{2}$ (对称伯努利随机变量), $S_n=\xi_1+\cdots+\xi_n$, $P_n(k)=\mathbf{P}\{S_n=k\}$, 其中 $k\in E_n=\{0,\pm1,\cdots,\pm n\}$.

利用全概率公式 (第 3 节中的 (3) 式), 证明如下的递推关系式成立:

$$P_{n+1}(k)=\frac{1}{2}P_n(k+1)+\frac{1}{2}P_n(k-1),\quad k\in E_{n+1},\tag{*}$$

或等价地, 有

$$P_{n+1}(k)-P_n(k)=\frac{1}{2}\Big(P_n(k+1)-2P_n(k)+P_n(k-1)\Big).\tag{**}$$

5. (续第 4 题) 随机变量序列 $S_0=0$, $S_1=\xi_1$, $S_2=\xi_1+\xi_2,\cdots$, $S_n=\xi_1+\cdots+\xi_n$ 可以看成是质点作随机游动的轨迹, 它自原点出发, 在每个整点时刻或者向上游动一个单位, 或者向下游动一个单位.

从现在起, 我们假定游动发生在时刻 $\Delta,2\Delta,\cdots,n\Delta$, 其中 $\Delta>0$, 并且在每个这样的时刻, 质点或者向上游动量 Δx, 或者向下游动量 Δx. 取代上题中的 $P_n(k)=\mathbf{P}\{S_n=k\}$, 我们来考察概率

$$P_{n\Delta}(k\Delta x)=\mathbf{P}\Big\{S_{n\Delta}=k\Delta x\Big\}.$$

通过与递推关系式 (**) 类比, 我们得到

$$\frac{P_{(n+1)\Delta}(k\Delta x)-P_{n\Delta}(k\Delta x)}{\Delta}=\frac{1}{2}\Big(P_{n\Delta}((k+1)\Delta x)-2P_{n\Delta}(k\Delta x)+P_{n\Delta}((k-1)\Delta x)\Big).$$

亦即 “关于时间的一阶导数的离散形式” 完全等同于 “关于空间位移的二阶导数离散形式” 的 $\frac{1}{2}$ 倍.

我们设 $\Delta x=\sqrt{\Delta}$, 并假定对于 $t>0$, $x\in\mathbb{R}$, 在 $n\to\infty$, $k\to\infty$ 时, 有 $n\Delta\to t$, $k\sqrt{\Delta}\to x$. 证明, 在这样的极限过程中, 有

(a) $P_t(x)=\lim P_{n\Delta}(k\sqrt{\Delta})$.

(b) 极限函数 $P_t(x)$ 满足热传导 “微分” 方程:

$$\frac{\partial P_t(x)}{\partial t}=\frac{1}{2}\frac{\partial^2 P_t(x)}{\partial x^2}$$

(贝塞尔, 爱因斯坦).

6. 作为上一题的推广, 设作随机游动的质点每一步以概率 $p(\Delta)=1/2+\Delta x$ 往上游动距离 Δx, 以概率 $q(\Delta)=1/2-\Delta x$ 往下游动距离 Δx. 仍设 $\Delta x=\sqrt{\Delta}$, $n\Delta\to t$, $k\sqrt{\Delta}\to x$. 证明, 如同上一题一样, 存在极限 $P_t(x)=\lim P_{n\Delta}(k\sqrt{\Delta})$, 并且 $P_t(x)$ 满足方程

$$\frac{\partial P_t(x)}{\partial t}=-2\frac{\partial P_t(x)}{\partial x}+\frac{1}{2}\frac{\partial^2 P_t(x)}{\partial x^2}.$$

7. 应当对上面两题中的极限过程作怎样的改变, 才能使得极限函数满足如下的方程

$$\frac{\partial P_t(x)}{\partial t}=-\mu\frac{\partial P_t(x)}{\partial x}+\frac{1}{2}\sigma^2\frac{\partial^2 P_t(x)}{\partial x^2}?$$

(该方程称为福克尔–普朗克方程, 或柯尔莫戈洛夫前向方程.)

8. 设 $F_n = F_n(t)$, $t \in [0,1]$, $n \geqslant 1$, 为非降函数序列, 对任何有理数 $t \in \mathbb{Q} \cap [0,1]$, 都有 $F_n(t) \to t$. 证明, 这种收敛性是一致的 (参阅第三章第 1 节第 5 题), 即

$$\sup_{t\in[0,1]} |F_n(t) - t| \to 0, \quad n \to \infty.$$

9. 证明, 如下的不等式对一切 $x > 0$ 都成立:

$$\frac{x}{1+x^2}\varphi(x) < 1 - \Phi(x) < \frac{\varphi(x)}{x},$$

其中 $\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$, $\Phi(x) = \int_{-\infty}^{x} \varphi(y)dy$.

提示: 考察导数 $\varphi'(x)$ 和 $(x^{-1}\varphi(x))'$.

10. 证明, 对于泊松分布, 有如下的局部定理成立: 对任何 $k = 0, 1, 2, \cdots$, 当 $\lambda \to \infty$ 时, 都有

$$\sqrt{\lambda}\left|\frac{\lambda^k}{k!}e^{-\lambda} - \frac{1}{\sqrt{2\pi\lambda}}\exp\left\{-\frac{1}{2\lambda}(k-\lambda)^2\right\}\right| \to 0.$$

提示: 运用斯特林公式.

§7. 伯努利概型中 "成功" 概率的估计

1. 假设先验地知道, 参数 θ 取值于集合 $\Theta_0 \subseteq [0,1]$. 试说明, 何时存在对于仅仅取值于该集合 Θ_0 的 θ 的无偏估计?

提示: 如果 Θ_0 是单点集 ($\Theta_0 = \{\theta_0\}$), 那么值 θ_0 就是所要求的估计. 如果 Θ_0 至少含有两个元素, 那么为了存在无偏估计, 必须且只需, $\{0\} \in \Theta_0$ 和 $\{1\} \in \Theta_0$. 试证之.

2. 试在上题中的条件下, 寻找相应的拉奥–克拉默不等式, 并考察有效估计.

3. 试在第 1 题中的条件下, 考察关于 θ 的置信区间的构造问题.

4. 在第 2 节第 21 题中进一步假定 N 足够大, $N \gg M$, $N \gg n$, 试考察估计 $\widehat{N}$ 的无偏性和有效性. 试仿照关于参数 θ 的置信区间的构造方式 (参阅《概率》第一卷第 74 页中的 (8) 式和 (9) 式), 构造关于 N 的具有如下性质的置信区间 $[\widehat{N} - a(\widehat{N}), \widehat{N} + b(\widehat{N})]$:

$$\mathbf{P}_{N,M;n}\{\widehat{N} - a(\widehat{N}) \leqslant N \leqslant \widehat{N} + b(\widehat{N})\} \approx 1 - \varepsilon,$$

其中 ε 是某个小的正数.

5. (χ^2 相合性准则) 设 $\xi_1, \cdots, \xi_n$ 为相互独立的伯努利随机变量, 有 $\mathbf{P}\{\xi_i = 1\} = p$, $\mathbf{P}\{\xi_i = 0\} = 1 - p$, $1 \leqslant i \leqslant n$. 与第 7 节中主要研究关于 "成功" 概率 p 的估计不同, 我们这里讨论利用观察结果 $x = (x_1, \cdots, x_n)$ 验证假设 H_0: $p = p_0$ 的问题, 亦即 p 的真实值是否等于给定的值 $0 < p_0 < 1$ 的问题.

记 $S_n(\xi) = \xi_1 + \cdots + \xi_n$, 令

$$\chi_n^2(\xi) = \frac{(S_n(\xi) - np_0)^2}{np_0(1-p_0)}.$$

试在假设 H_0 成立的假定之下, 证明, 对一切 $x \geqslant 0$, 都有

$$\mathbf{P}\left\{\chi_n^2(\xi) \leqslant x\right\} \to \int_0^x \frac{1}{\sqrt{2\pi y}}e^{-y/2}dy, \quad n \to \infty.$$

(根据第二章第 3 节中的表 3, 函数 $F(x) = \int_0^x \frac{1}{\sqrt{2\pi y}}e^{-y/2}dy$ 是自由度为 1 的随机变量 χ^2 的分布函数, 亦即参数为 0 和 1 的高斯随机变量之平方的分布函数.)

在检验假设 H_0: $p = p_0$ 的 χ^2 相合性准则的基础上, 产生出如下的讨论模式:

将 $\varepsilon > 0$ 取得如此之小, 使得在一次试验中, 具有概率 ε 的事件的发生非常罕见. (例如取 $\varepsilon = 0.01$, 那么根据第 5 节中关于 (8) 式的注释, 在 100 次试验中, 具有概率 0.01 的事件平均起来, 只会发生一次.)

根据所取的 $\varepsilon > 0$, 求出使得等式 $\int_{\lambda(\varepsilon)}^{\infty} \frac{1}{\sqrt{2\pi y}}e^{-y/2}dy = \varepsilon$ 成立的 $\lambda(\varepsilon)$.

(根据 χ^2 相合性准则) 按照如下方式检验假设 H_0: $p = p_0$, 如果根据观察值 $x = (x_1, \cdots, x_n)$ 计算出的量 $\chi_n^2(x)$ 超过 $\lambda(\varepsilon)$, 那么就拒绝假设 H_0: $p = p_0$, 而如果假设 $\chi_n^2(x) \leqslant \lambda(\varepsilon)$, 则接受假设 H_0, 亦即认为观察值 $x = (x_1, \cdots, x_n)$ 与假设 $p = p_0$ 相合.

(a) 试利用大数定律 (参阅第 5 节) 验证断言: 根据 χ^2 相合性准则对假设 H_0: $p = p_0$ 所做的相合性检验的结论 (至少在 n 较大时) 是唯一确定的.

(b) 试利用第 6 节 (24) 式中的贝里–埃森 (Berry-Esseen) 不等式, (在 $p = p_0$ 的假设之下) 证明,

$$\sup_x \left|\mathbf{P}\left\{\chi_n^2(\xi) \leqslant x\right\} - \int_0^x \frac{1}{\sqrt{2\pi y}}e^{-y/2}dy\right| \leqslant \frac{2}{\sqrt{np_0(1-p_0)}}.$$

(c) 设 $\lambda_n(\varepsilon)$ 由不等式 $\mathbf{P}\left\{\chi_n^2(\xi) \geqslant \lambda_n(\varepsilon)\right\} \leqslant \varepsilon$ 所确定. 试求 $\lambda_n(\varepsilon)$ 趋于 $\lambda(\varepsilon)$ 的收敛速度, 亦即设法确定在运用 χ^2 相合性准则时, 以事件 "$\chi_n^2(\xi) \geqslant \lambda_n(\varepsilon)$" 取代事件 "$\chi_n^2(\xi) \geqslant \lambda(\varepsilon)$" 时所产生的误差.

6. 设 ξ 为随机变量, 服从二项分布, 有

$$\mathbf{P}_\theta\{\xi = k\} = C_n^k \theta^k (1-\theta)^{n-k}, \quad 0 \leqslant k \leqslant n,$$

其中 n 为给定值, 而 θ 为 "未知参数", 要求根据随机变量 ξ 的 (唯一) 观察值估计 θ.

θ 的标准估计是 $T(\xi) = \xi/n$, 它具有无偏性: 对一切 $\theta \in [0,1]$, 都有

$$\mathbf{E}_\theta T(\xi) = \theta.$$

证明, 在无偏估计类 $\widetilde{T}=\widetilde{T}(\xi)$ 中, 该估计是有效的, 即

$$\mathbf{E}_\theta(T(\xi)-\theta)^2=\inf_{\widetilde{T}}\mathbf{E}_\theta(\widetilde{T}(\xi)-\theta)^2.$$

再证明, 如果 $n=3$, 并且先验知道 $\theta\in(1/4,3/4)$, 则对于任何 $\theta\neq 1/2$, 估计 $\widehat{T}(\theta)=1/2$ 都是有偏的, 但是却比无偏估计 $T(\xi)=\xi/3$ 更好:

$$\mathbf{E}_\theta(\widehat{T}(\xi)-\theta)^2<\mathbf{E}_\theta(T(\xi)-\theta)^2,$$

亦即

$$\mathbf{E}_\theta\left(\frac{1}{2}-\theta\right)^2<\mathbf{E}_\theta\left(\frac{\xi}{3}-\theta\right)^2.$$

试对任何 n 讨论类似断言的正确性.

7. 两位校对员 A 和 B 分别发现了 a 处和 b 处印刷错误, 其中有 c 处是共同的. 假定两位校对员相互独立地工作, 试对尚未发现的印刷错误的数目给出一个 “合乎情理的” 估计.

提示: 假定印刷错误的总数 n 很大, 而 a/n 与 b/n 是关于两位校对员 A 和 B 分别发现一处错误的概率 p_a 和 p_b 的足够好的估计, 再进行有关概率的讨论.

§8. 关于分割的条件概率与条件数学期望

1. 试给出两个随机变量 ξ 和 η, 它们不相互独立, 但是却有

$$\mathbf{E}(\xi|\eta)=\mathbf{E}\xi.$$

(试与命题 (22) 相比较. 此处以及下面, 均适用本习题集本章第 4 节开头的注释.)

2. 随机变量 ξ 关于分割 $\mathscr{D}$ 的*条件方差*定义为如下的随机变量:

$$\mathbf{D}(\xi|\mathscr{D})=\mathbf{E}\Big((\xi-\mathbf{E}(\xi|\mathscr{D}))^2|\mathscr{D}\Big).$$

证明

$$\mathbf{D}\xi=\mathbf{E}\mathbf{D}(\xi|\mathscr{D})+\mathbf{D}\mathbf{E}(\xi|\mathscr{D}).$$

提示: 验证

$$\mathbf{E}\mathbf{D}(\xi|\mathscr{D})=\mathbf{E}\xi^2-\mathbf{E}\big(\mathbf{E}(\xi|\mathscr{D})\big)^2 \quad 和 \quad \mathbf{D}\mathbf{E}(\xi|\mathscr{D})=\mathbf{E}\big(\mathbf{E}(\xi|\mathscr{D})\big)^2-(\mathbf{E}\xi)^2.$$

3. 试从 (17) 式出发, 证明, 对任何函数 $f=f(\eta)$, 条件数学期望 $\mathbf{E}(\xi|\eta)$ 都具有如下性质

$$\mathbf{E}\big(f(\eta)\mathbf{E}(\xi|\eta)\big)=\mathbf{E}\big(\xi f(\eta)\big).$$

4. 设 ξ 与 η 为随机变量. 证明, $\inf_f\mathbf{E}(\eta-f(\xi))^2$ 在函数 $f^*(\xi)=\mathbf{E}(\eta|\xi)$ 处达到. (据此结果, 根据 ξ 对 η 所作的最小方差意义下的估计就是 $\mathbf{E}(\eta|\xi)$.)

提示: 试验证对任意函数 $f=f(x)$, 都有

$$\mathbf{E}(\eta-f(\xi))^2=\mathbf{E}(\eta-f^*(\xi))^2+2\mathbf{E}\big[(\eta-f^*(\xi))(f^*(\xi)-f(\xi))\big]+\mathbf{E}(f^*(\xi)-f(\xi))^2,$$

其中中括号 $[\cdot]$ 中的随机变量的数学期望为 0.

5. 设 $\xi_1,\cdots,\xi_n,\tau$ 为相互独立的随机变量, 并且 $\xi_1,\cdots,\xi_n$ 同分布, 而 τ 取值 $1,\cdots,n$, 令 $S_\tau=\xi_1+\cdots+\xi_\tau$ 为随机个随机变量的和. 证明

$$\mathbf{E}(S_\tau|\tau)=\tau\mathbf{E}\xi_1,\quad \mathbf{D}(S_\tau|\tau)=\tau\mathbf{D}\xi_1,$$
$$\mathbf{E}S_\tau=\mathbf{E}\tau\cdot\mathbf{E}\xi_1,\quad \mathbf{D}S_\tau=\mathbf{E}\tau\cdot\mathbf{D}\xi_1+\mathbf{D}\tau\cdot(\mathbf{E}\xi_1)^2.$$

提示: 在证明后两个等式时, 可利用前两个等式, 即

$$\mathbf{E}(S_\tau|\tau)=\tau\mathbf{E}\xi_1 \quad 和 \quad \mathbf{D}(S_\tau|\tau)=\tau\mathbf{D}\xi_1.$$

6. 证明等式 (24).①

7. 设 $\mathscr{E}$ 是具有基本结果空间 $\Omega=\{\omega_1,\cdots,\omega_k\}$ 和概率 (“权重”) $p_i=p(\omega_i)$, $\sum_{i=1}^k p_i=1$ 的试验. 在第 5 节 (14) 式中曾将 $H=-\sum_{i=1}^k p_i\ln p_i$ 定义为分布 $(p_1,\cdots,p_k)$ 的熵, 并以其作为试验 $\mathscr{E}$ 的 “不确定性” 程度的度量. 在那里还证明了, “不确定性” 的最大值在所有 k 个不同结果均以相同的概率出现的试验中达到, 此时 $H=\ln k$.

对数函数 (在等概场合下) 作为 “不确定性” 程度的度量是很自然的, 其自然性由如下命题 (也是本题所要求证明的结论) 所确定:

假设具有 k 个基本结果的试验 $\mathscr{E}$ 的 “不确定性” 程度由具有下述性质的函数 $f(k)$ 来确定: $f(1)=0$, 而若 $k>l$, 则 $f(k)>f(l)$. 并且我们假设 $f(kl)=f(k)+f(l)$. (这一性质反映了这样一种需求: 对于分别具有 k 和 l 个基本结果的相互独立的试验 $\mathscr{E}_1$ 和 $\mathscr{E}_2$, 如果我们将试验 $\mathscr{E}_1\bigotimes\mathscr{E}_2$ 定义为同时进行两个试验 $\mathscr{E}_1$ 和 $\mathscr{E}_2$, 那么 $\mathscr{E}_1\bigotimes\mathscr{E}_2$ 的 “不确定性” 程度就应当是两个试验的 “不确定性” 程度之和.)

证明, 在所述的条件之下, 函数 $f(k)$ 具有形式 $f(k)=c\log_b k$, 其中 $c>0$ 为常数, 而 $\log_b k$ 是以 b 为底的对数.

① 此处所要求证明的 (24) 式为: 设随机变量 ξ 与分割 $\mathscr{D}$ 独立, 即对任何 $D_i\in\mathscr{D}$, 随机变量 ξ 都与随机变量 I_{D_i} 相互独立. 证明

$$\mathbf{E}(\xi\,|\,\mathscr{D})=\mathbf{E}\xi$$

—— 译者注.

注: 由于在不同底的对数之间存在换底公式 $\log_b k = \log_b a \cdot \log_a k$, 所以当以不同的底的对数度量 "不确定性" 程度时, 只不过是所采用的计量单位不同罢了. 通常人们采用 $b=2$. 由于对于 $k=2$, 有 $\log_2 k = 1$, 因此可以说, 在具有两个等概的基本结果的试验中, "不确定性" 的程度等于 1 (个单位). 在通信理论 (尤其是在编码理论) 中, 这种 "不确定性" 单位称为二进制单位或 BIT(英文词汇 BInary digiT 的缩写). 如此一来, 如果试验 $\mathscr{E}$ 是具有 $k=10$ 个等概的基本结果, 那么它的 "不确定性" 的程度就等于 $\log_2 10 (\approx 3.3)$BIT.

8. 设 $(\Omega, \mathscr{F}, \mathbf{P})$ 是某个离散概率空间, $\xi = \xi(\omega)$ 是随机变量, 它分别以概率 $\mathbf{P}\{\xi = x_i\} = p_i$ 取值 $x_1, \cdots, x_k$. 我们将

$$H(\xi) = -\sum_{i=1}^{k} p_i \log_2 p_i$$

称为随机变量 ξ 的熵 (或由 ξ 的观察值构成的试验 $\mathscr{E}_\xi$ 的熵).(试与第 5 节 (14) 式相比较, 那里不是使用以 2 为底的对数 $\log_2$, 而是采用自然对数 ln. 正如我们在上题中所阐明的那样, 它们之间没有原则性的区别.)

如果 (ξ, η) 为随机变量对, 有 $\mathbf{P}\{\xi = x_i, \eta = y_j\} = p_{ij}$, $i = 1, \cdots, k, j = 1, \cdots, l$, 则该对随机变量的熵 $H(\xi, \eta)$ 相应地定义为

$$H(\xi, \eta) = -\sum_{i=1}^{k} \sum_{j=1}^{l} p_{ij} \log_2 p_{ij}.$$

证明, 如果 ξ 与 η 相互独立, 则 $H(\xi, \eta) = H(\xi) + H(\eta)$.

9. 设 (ξ, η) 为随机变量对, 它们的所可能取的值为 $(x_i,\ y_j)$, $i = 1, \cdots, k, j = 1, \cdots, l$. 我们将

$$H_{x_i}(\eta) = -\sum_{j=1}^{l} \mathbf{P}\{\eta = y_j \mid \xi = x_i\} \log_2 \mathbf{P}\{\eta = y_j \mid \xi = x_i\}$$

称为随机变量 η 关于事件 $\{\xi = x_i\}$ 的条件熵. 而 η 关于 ξ 的平均条件熵则定义为

$$H_\xi(\eta) = \sum_{i=1}^{k} \mathbf{P}\{\xi = x_i\} H_{x_i}(\eta).$$

证明

(a) $H(\xi, \eta) = H(\xi) + H_\xi(\eta)$.

(b) 如果 ξ 与 η 相互独立, 则 $H(\xi, \eta) = H(\xi) + H(\eta)$.

(c) $0 \leqslant H_\xi(\eta) \leqslant H(\eta)$.

10. 设 (ξ, η) 为随机变量对, 称

为含于 ξ 中的关于 η 的信息量. 该术语的含义是: 差值 $H(\eta) - H_\xi(\eta)$ 表明随机变量 ξ 可以在多大程度上减小随机变量 η 的 "不确定性".

证明

(a) $I_\xi(\eta) = I_\eta(\xi) \geqslant 0$.

(b) $I_\xi(\eta) = H(\eta)$ 当且仅当 η 是 ξ 的某个函数.

(c) 如果 ξ, η 和 ζ 为随机变量, 则

$$I_{(\xi,\zeta)}(\eta) = H(\eta) - H_{(\xi,\zeta)}(\eta) \geqslant I_\xi(\eta),$$

意即随机变量对 (ξ, ζ) 中所包含的关于 η 的信息不少于单个随机变量 ξ 中所包含的信息.

11. 设 $\xi_1, \cdots, \xi_n$ 为独立同分布的伯努利随机变量, 有 $\mathbf{P}\{\xi_1 = 1\} = p$, $\mathbf{P}\{\xi_1 = 0\} = 1 - p$, 令 $S_n = \xi_1 + \cdots + \xi_n$. 证明

(a) $\mathbf{P}\{\xi_1 = x_1, \cdots, \xi_n = x_n \mid S_n = k\} = \dfrac{I_{\{x\}}(k)}{C_n^x}$.

(b) $\mathbf{P}\{S_n = x \mid S_{n+m} = k\} = \dfrac{C_n^x C_m^{k-x}}{C_{n+m}^k}$,

其中 $x = x_1 + \cdots + x_n$, $x_i = 0, 1$, 而且 $x \leqslant k$.

§9. 随机游动 I. 掷硬币博弈的破产概率和平均持续时间

1. 证明, 作为 (33) 和 (34) 式推广, 有

$$\mathbf{E} S_{\tau_n^x}^x = x + (p - q)\mathbf{E}\tau_n^x,$$
$$\mathbf{E}\left(S_{\tau_n^x}^x - \tau_n^x \mathbf{E}\xi_1\right)^2 = \mathbf{D}\xi_1 \cdot \mathbf{E}\tau_n^x + x^2.$$

2. 试研究当水平 $A \downarrow -\infty$ 时, 量 $\alpha(x)$, $\beta(x)$, $m(x)$(定义见《概率》第一卷第 85 页的 (13) 式和第 89 页的 (26) 式) 分别收敛于何处?

提示: 答案如下:

$$\lim_{A \to -\infty} \alpha(x) = \begin{cases} 0, & p \geqslant q, \\ \dfrac{(q/p)^\beta - (q/p)^x}{(q/p)^\beta}, & p < q. \end{cases}$$

$$\lim_{A \to -\infty} m(x) = \begin{cases} \dfrac{\beta - x}{p - q}, & p > q, \end{cases}$$

3. 设在独立同分布伯努利随机变量序列中, 有 $p=q=1/2$. 证明

$$\mathbf{E}|S_n| \sim \sqrt{\frac{2}{\pi}n}, \quad n\to\infty. \tag{*}$$

提示: 可以直接推出如下的关系式 (即所谓离散型"田中公式 (Tannaka formula)", 亦可参阅第七章第 9 节第 8 题): 对 $n\geqslant 1$, 有

$$|S_n| = \sum_{k=1}^{n} \operatorname{sign}(S_{k-1})\Delta S_k + N_n, \tag{**}$$

其中 $S_0=0,\ S_k=\xi_1+\cdots+\xi_k,\ \Delta S_k=\xi_k$,

$$\operatorname{sign} x = \begin{cases} 1, & x>0,\\ 0, & x=0,\\ -1, & x<0,\end{cases}$$

而 $N_n=\#\{0\leqslant k\leqslant n-1:\ S_k=0\}$, 即使得 $S_k=0$ 的 $k\,(0\leqslant k\leqslant n-1)$ 的个数. 于是就有

$$\mathbf{E}|S_n| = \mathbf{E}N_n = \mathbf{E}\sum_{k=1}^{n-1} I(S_k=0) = \sum_{k=1}^{n-1}\mathbf{P}\{S_k=0\}. \tag{***}$$

再利用 $\mathbf{P}\{S_{2k}=0\}=2^{-2k}C_{2k}^{k}$, $\mathbf{P}\{S_k=0\}=0$.

注: 由 (***) 可推知

$$\mathbf{E}N_n \sim \sqrt{\frac{2}{\pi}n}, \quad n\to\infty.$$

(亦可比较《概率》第二卷第七章第 9 节例 2, 不过应将该例中 (17) 和 (18) 式中的 $1/2\pi$ 换为 $2/\pi$.)

4. 两名游戏者相互独立地抛掷自己手中的均匀硬币 (每人一枚). 证明, 他们在各自抛掷的 n 次之中所抛出的正面次数相等的概率为 $2^{-2n}\sum\limits_{k=0}^{n}(C_n^k)^2$. 并由此推出等式 $\sum\limits_{k=0}^{n}(C_n^k)^2=C_{2n}^{n}$ (亦可参阅第 2 节第 2 题).

设 σ_n 为首次出现两人的正面相等的时刻 (如果直到掷完 n 次, 都未出现两人的正面相等, 则令 $\sigma_n=n+1$). 试求概率 $\mathbf{P}\{\sigma_n=k\}$, $1\leqslant k\leqslant n+1$, 并求数学期望 $\mathbf{E}\min(\sigma_n,\,n)$.

提示: 如果第 k 个人在第 i 次抛掷时掷出正面, 就令 $\xi_i^{(k)}=1$, 掷出反面, 就令 $\xi_i^{(k)}=-1$, $k=1,2$. 那么就有

$$\begin{aligned}\mathbf{P}\{\text{两人在 } n \text{ 次抛掷中掷出的正面次数相等}\} &= \mathbf{P}\left\{\sum_{i=1}^{n}\xi_i^{(1)}=\sum_{i=1}^{n}\xi_i^{(2)}\right\}\\ &= \sum_{j=0}^{n}\mathbf{P}\left\{\sum_{i=1}^{n}\xi_i^{(1)}=2j-n,\ \sum_{i=1}^{n}\xi_i^{(2)}=2j-n\right\}=\sum_{j=0}^{n}2^{-2n}(C_n^j)^2\end{aligned}$$

以及

$$\mathbf{P}\left\{\sum_{i=1}^{n}\xi_i^{(1)}=\sum_{i=1}^{n}\xi_i^{(2)}\right\}=\mathbf{P}\left\{\sum_{i=1}^{2n}\eta_i=0\right\}=2^{-2n}C_{2n}^{n},$$

其中 $\eta_1=\xi_1^{(1)},\ \eta_2=-\xi_1^{(2)},\ \eta_3=\xi_2^{(1)},\ \eta_4=-\xi_2^{(2)},\ \cdots$.

5. 设 $\xi_1,\cdots,\xi_N$ 为相互独立的对称伯努利随机变量, 即有 $\mathbf{P}\{\xi_i=1\}=\mathbf{P}\{\xi_i=-1\}=1/2$, 记 $S_n=\xi_1+\cdots+\xi_n$, $1\leqslant n\leqslant N$. 试求概率

$$\mathbf{P}\left(\bigcup_{N_1<n\leqslant N_2}\{S_n=0\}\right),$$

亦即在集合 $(N_1+1,\cdots,N_2)$ 中存在 n, 使得 $S_n=0$ 的概率, 其中 $1\leqslant N_1+1\leqslant N_2\leqslant N$.

6. 设 $\xi_0,\xi_1,\cdots,\xi_N$ 为相互独立的对称伯努利随机变量, 即有 $\mathbf{P}\{\xi_i=1\}=\mathbf{P}\{\xi_i=-1\}=1/2$, 记 $S_n=\xi_1+\cdots+\xi_n$ 和 $X_n=\xi_0(-1)^{S_n}$ 为离散的电报信号, $1\leqslant n\leqslant N$. 试求随机变量 X_n 的分布与方差①.

并求条件概率

$$\mathbf{P}\{X_n=1 \mid \xi_0=i\}, \quad i=\pm1, \quad 1\leqslant n\leqslant N.$$

7. 设 $\xi_1,\cdots,\xi_N$ 为独立同分布的伯努利随机变量, 有 $\mathbf{P}\{\xi_i=1\}=p$, $\mathbf{P}\{\xi_i=-1\}=1-p$. 记 $S_i=\xi_1+\cdots+\xi_i$, $1\leqslant i\leqslant N$, $S_0=0$. 设 $\mathscr{R}_N$ 为随机游动 $S_0,S_1,\cdots,S_N$ 的变程, 即它们的值集中的不同元素的个数.

试求 $\mathbf{E}\mathscr{R}_N$ 和 $\mathbf{D}\mathscr{R}_N$. 试问, 对于怎样的 p, 变程 $\mathscr{R}_N$ 满足如下形式的大数定律

$$\mathbf{P}\left\{\left|\frac{\mathscr{R}_N}{N}-c\right|>\varepsilon\right\}\to 0, \quad N\to\infty?$$

其中 $\varepsilon>0$, 而 c 为某个常数. (亦可参阅第二章第 6 节第 87 题和第八章第 8 节第 16 题.)

① 原文为: 求 X_n 的值与方差 —— 译者注.

8. 设 $\xi_1, \cdots, \xi_N$ 为同分布随机变量 (不一定为伯努利随机变量)[①], 记 $S_0 = 0$, $S_i = \xi_1 + \cdots + \xi_i$, $1 \leqslant i \leqslant N$.

令

$$N_n = \sum_{k=1}^{n} I(S_k > 0),$$

即序列 $S_0, S_1, \cdots, S_n$ 中的正数的个数. 证明如下结果 (斯帕尔–安德森)

$$\mathbf{P}\{N_n = k\} = \mathbf{P}\{N_k = k\}\mathbf{P}\{N_{n-k} = 0\}, \quad 0 \leqslant k \leqslant n.$$

9. 设 $\xi_1, \cdots, \xi_N$ 为第 7 题中的伯努利随机变量, 我们按如下方式定义变量 $X_1, \cdots, X_N$, 其中

$$X_1 = \xi_1, \quad X_n = \lambda X_{n-1} + \xi_n, \quad 2 \leqslant n \leqslant N, \ \lambda \in \mathbb{R}.$$

试求 $\mathbf{E}X_n, \mathbf{D}X_n$ 和 $\mathbf{cov}(X_n, \ X_{n+k})$.

§10. 随机游动 II. 反射原理. 反正弦定律

1. 当 $n \to \infty$ 时, $\mathbf{E}\min(\sigma_{2n}, \ 2n)$ 以怎样的速度趋于 ∞? (其中 $\sigma_{2n} = \min\{1 \leqslant k \leqslant 2n: \ S_k = 0\}$; 而如果对一切 $1 \leqslant k \leqslant 2n$, 都有 $S_k \neq 0$, 则令 $\sigma_{2n} = \infty$ 或 $\sigma_{2n} = 2n$.)

提示: 根据命题 1, 我们有

$$\mathbf{E}\min(\sigma_{2n}, \ 2n) = \sum_{k=1}^{n} u_{2(k-1)} + 2nu_{2n},$$

其中 $u_{2n} \sim 1/\sqrt{\pi n}$. 因此可知

$$\mathbf{E}\min(\sigma_{2n}, \ 2n) \sim 4\sqrt{\frac{n}{\pi}}, \quad n \to \infty.$$

2. 设 $\tau_n = \min\{1 \leqslant k \leqslant n: \ S_k = 1\}$, 而若对一切 $1 \leqslant k \leqslant n$, 都有 $S_k < 1$, 则令 $\tau_n = \infty$. 试问: 对于对称的伯努利游动 $(p = q = 1/2)$ 和不对称的伯努利游动 $(p \neq q)$, 当 $n \to \infty$ 时, $\mathbf{E}\min(\tau_n, n)$ 分别趋向于何处?

提示: 本题答案为

$$\mathbf{E}\min(\tau_n, n) \to \begin{cases} (p-q)^{-1}, & p > q, \\ \infty, & p \leqslant q. \end{cases}$$

①原文中就没有独立性假设 —— 译者注.

3. 以《概率》第一卷第 10 节中的思想与方法, 证明, 对于对称的伯努利随机游动 $(p = q = 1/2)$, 如果记 $S_0 = 0$, $S_k = \xi_1 + \cdots + \xi_k$, $k \leqslant n$, 则有如下各关系式成立 (N 为正整数):

$$\mathbf{P}\Big\{\max_{1 \leqslant k \leqslant n} S_k \geqslant N, \ S_n < N\Big\} = \mathbf{P}\{S_n > N\},$$

$$\mathbf{P}\Big\{\max_{1 \leqslant k \leqslant n} S_k \geqslant N\Big\} = 2\mathbf{P}\{S_n \geqslant N\} - \mathbf{P}\{S_n = N\},$$

$$\mathbf{P}\Big\{\max_{1 \leqslant k \leqslant n} S_k = N\Big\} = \mathbf{P}\{S_n = N\} + \mathbf{P}\{S_n = N+1\} = 2^{-n}C_n^{\left[\frac{n+N+1}{2}\right]},$$

$$\mathbf{P}\Big\{\max_{1 \leqslant k \leqslant n} S_k \leqslant 0\Big\} = \mathbf{P}\{S_n = 0\} + \mathbf{P}\{S_n = 1\} = 2^{-n}C_n^{\left[\frac{n}{2}\right]},$$

$$\mathbf{P}\Big\{\max_{1 \leqslant k \leqslant n-1} S_k \leqslant 0, \ S_n > 0\Big\} = \mathbf{P}\{S_1 \neq 0, \ \cdots, \ S_n \neq 0, \ S_{n+1} = 0\},$$

$$\mathbf{P}\{S_1 > 0, \ \cdots, \ S_{2n-1} > 0, \ S_{2n} = 0\} = \frac{1}{n}2^{-2n}C_{2n-2}^{n-1},$$

$$\mathbf{P}\{S_1 \geqslant 0, \ \cdots, \ S_{2n-1} \geqslant 0, \ S_{2n} = 0\} = \frac{1}{n+1}2^{-2n}C_{2n}^{n}.$$

再证明

$$\mathbf{P}\{S_{2n} = 2k\} = 2^{-2n}C_{2n}^{n-k}, \quad k = 0, \pm 1, \cdots, \pm n,$$

与

$$\mathbf{P}\{S_{2n+1} = 2k+1\} = 2^{-2n-1}C_{2n+1}^{n-k}, \quad k = -(n+1), 0, \pm 1, \cdots, \pm n,$$

上述二式可以统一地写为

$$\mathbf{P}\{S_n = k\} = \begin{cases} 2^{-n}C_n^{(n-k)/2}, & \text{如果} k \equiv n \ (\mathrm{mod}\ 2), \\ 0, & \text{其他情况}, \end{cases}$$

其中 $k = 0, \pm 1, \cdots, \pm n$.

作为上面仅仅关于正整数 N 的概率关系式 $\mathbf{P}\Big\{\max\limits_{1 \leqslant k \leqslant n} S_k = N\Big\}$ 的补充, 对于 $r = 0, 1, \cdots, n$, 证明

$$\mathbf{P}\Big\{\max_{0 \leqslant k \leqslant n} S_k = r\Big\} = 2^{-n}C_n^{[(n-r)/2]}.$$

4. 设 $\xi_1, \cdots, \xi_{2n}$ 为独立同分布伯努利随机变量, 有 $\mathbf{P}\{\xi_k = 1\} = \mathbf{P}\{\xi_k = -1\} = 1/2$, $k \leqslant 2n$. 令 $S_0 = 0$, $S_k = \xi_1 + \cdots + \xi_k$, 而

$$g_{2n} = \max\{0 < 2k \leqslant 2n: \ S_{2k} = 0\},$$

即序列 $S_2, S_4, \cdots, S_{2n}$ 中最后一个 0 的出现时刻, 若这种时刻不存在, 就令 $g_{2n} = 0$.

证明

$$\mathbf{P}\{g_{2n}=2k\}=u_{2n}u_{2(n-k)},\quad 1\leqslant k\leqslant n,$$

其中 $u_{2k}=\mathbf{P}\{S_{2k}=0\}=2^{-2k}C_{2k}^{k}$.

(以 $P_{2k,2n}$ 表示质点在区间 $[0,2n]$ 中有 $2k$ 次向右移动的概率 (参阅第 10 节 (12) 式), 试将 g_{2n} 的分布与 $P_{2k,2n}$ 相比较, 即可看出, 就像 (15) 式所表明的那样, 对于 $0<x<1$, 有

$$\sum_{\{k:\,0<k/n\leqslant x\}}\mathbf{P}\{g_{2n}=2k\}\to\frac{2}{\pi}\arcsin\sqrt{x},\quad n\to\infty,$$

亦即最后一个零点出现时刻的概率分布渐近于反正弦律.)

5. 在上一题的条件下, 以 θ_{2n} 表示 $S_0,S_1,\cdots,S_{2n}$ 中的最大值的首次出现时刻, 即如果 $S_0<S_k,\cdots,S_{k-1}<S_k$, 并且 $S_{k+1}\leqslant S_k,\cdots,S_{2n}\leqslant S_k$, 则有 $\theta_{2n}=k$; 而如果不存在这样的 $k>1$, 则令 $\theta_{2n}=0$.

证明

$$\mathbf{P}\{\theta_{2n}=0\}=u_{2n},\quad \mathbf{P}\{\theta_{2n}=2n\}=\frac{1}{2}u_{2n},$$

而对于 $0<k<n$, 则有

$$\mathbf{P}\{\theta_{2n}=2k \text{ 或 } 2k+1\}=\frac{1}{2}u_{2k}u_{2n-2k}.$$

并由此推出: (类似于上题) 对于最大值的首次出现时刻成立反正弦律: 对 $0<x<1$, 有

$$\sum_{\{k:\,0<k/n\leqslant x\}}\mathbf{P}\{\theta_{2n}=2k \text{ 或 } 2k+1\}\to\frac{2}{\pi}\arcsin\sqrt{x},\quad n\to\infty.$$

还可考察 $x=0$ 和 $x=1$ 的情形.

6. 设 $S_k=\xi_1+\cdots+\xi_k,\ k\leqslant 2n$, 其中 $\xi_1,\cdots,\xi_{2n}$ 为独立同分布随机变量, 有 $\mathbf{P}\{\xi_1=1\}=\mathbf{P}\{\xi_1=-1\}=1/2$. 证明

(a) 对于 $r=\pm1,\cdots,\pm n$, 有

$$\mathbf{P}\{S_1\neq0,\cdots,S_{2n-1}\neq0,S_{2n}=2r\}=C_{2n}^{n+r}\frac{|r|}{n}2^{-2n};$$

(b) 对于 $r=0,\pm1,\cdots,\pm n$, 有

$$\mathbf{P}\{S_{2n}=2r\}=C_{2n}^{n-r}2^{-2n}.$$

7. 设 $\xi_1,\cdots,\xi_n$ 为独立同分布随机变量, 有 $\mathbf{P}\{\xi_1=1\}=\mathbf{P}\{\xi_1=-1\}=1/2$. 记 $S_0=0,\ S_k=\xi_1+\cdots+\xi_k,\ 1\leqslant k\leqslant n$, 称 $\{S_k,\ k\leqslant n\}$ 为对称伯努利随机游动. 令

$$M_n=\max_{0\leqslant k\leqslant n}S_k,\quad m_n=\min_{0\leqslant k\leqslant n}S_k.$$

证明

$$(M_n-S_n,\ S_n-m_n,\ S_n)\overset{\text{law}}{=}(-m_n,\ M_n,\ S_n)\overset{\text{law}}{=}(M_n,\ -m_n,\ S_n),$$

其中 "$\overset{\text{law}}{=}$" 表示相应的三个随机变量的分布相同.

8. 设 $S_0=0,\ S_k=\xi_1+\cdots+\xi_k,\ k\geqslant1$, 其中 $\xi_1,\xi_2,\cdots$ 为独立随机变量, 有 $\mathbf{P}\{\xi_k=1\}=p,\ \mathbf{P}\{\xi_k=-1\}=q,\ p+q=1$. 证明

$$\mathbf{P}\Big\{\max_{1\leqslant k\leqslant n}S_k\geqslant N,\ S_n=m\Big\}=C_n^u p^v q^{n-v},$$

其中 $u=N+(n-m)/2$, 而 $v=(n+m)/2$, 并由此推出, 当 $p=q=1/2$ 和 $m\leqslant N$ 时, 有

$$\mathbf{P}\Big\{\max_{1\leqslant k\leqslant n}S_k=N,\ S_n=m\Big\}=\mathbf{P}\{S_n=2N-m\}-\mathbf{P}\{S_n=2N-m+2\}.$$

9. 设 $\xi_1,\xi_2,\cdots$ 为独立随机变量的无穷序列, 有 $\mathbf{P}\{\xi_i=1\}=\mathbf{P}\{\xi_i=-1\}=1/2$. 令 $S_0=0,\ S_n=\xi_1+\cdots+\xi_n$, 并且对于 $x\in\mathbb{Z}=\{0,\pm1,\pm2,\cdots\}$, 定义时刻 (时刻 0 之后首次到达状态 x 的时刻)

$$\sigma_1(x)=\inf\{n>0:\ \ S_n=x\},$$

如果 $\{\cdot\}=\varnothing$, 则令 $\sigma_1(x)=\infty$.

证明, 对于 $x=1,2,\cdots$, 有

$$\mathbf{P}\{\sigma_1(x)>n\}=\mathbf{P}\Big\{\max_{0\leqslant k\leqslant n}S_k<x\Big\},\quad \mathbf{P}\{\sigma_1(1)=2n+1\}=\frac{2^{-2n-1}}{n+1}C_{2n}^{n},$$

$$\mathbf{P}\{\sigma_1(x)=n\}=\frac{x}{n}2^{-n}C_n^{(n+x)/2},\quad \mathbf{P}\{\sigma_1(1)>n\}=2^{-n}C_n^{[n/2]}.$$

注: 关于独立随机变量无穷序列 $\xi_1,\xi_2,\cdots$ 的存在性问题, 见《概率》第一卷第一章第 5 节第 1 小节.

10. 设上题中的条件成立. 除了定义 $\sigma_1(x)$ 之外, 再定义时刻

$$\sigma_k(x)=\inf\{n>\sigma_{k-1}(x):\ \ S_n=x\},\quad k=2,3,\cdots,$$

如果 $\{\cdot\}=\varnothing$, 则令 $\sigma_k(x)=\infty$. (这一时刻的含义是清楚的, $\sigma_k(x)$ 就是第 k 次等于 x 的时刻.)

证明, 对于 $n=1,2,\cdots$, 有

$$\mathbf{P}\{\sigma_1(0)=2n\}=2^{-2n+1}n^{-1}C_{2n-2}^{n-1},\quad \mathbf{P}\{\sigma_1(0)<\infty\}=1,$$

$$\mathbf{P}\{\sigma_1(0)>2n\}=2^{-2n}C_{2n}^{n}=\mathbf{P}\{S_{2n}=0\},\quad \mathbf{E}\sigma_1(0)=\infty.$$

再证明, $\sigma_1(0),\sigma_2(0)-\sigma_1(0),\sigma_3(0)-\sigma_2(0),\cdots$ 形成独立同分布随机变量序列. (这一性质是 "再生周期" 方法的基础, 该方法被广泛地应用于随机游动的研究, 更详细的论述参阅本书附录第 7 节.)

11. 设 $\xi_1, \xi_2, \cdots$ 与上题中的相同. 令

$$L_n(x) = \#\{k,\ 0 < k \leqslant n:\ \ S_k = x\}.$$

($L_n(x)$ 就是随机游动 $\{S_k\}_{0<k\leqslant n}$ 到达状态 x 的时刻的数目, 其中 $x = 0, \pm 1, \pm 2, \cdots$; 试比较第七章第 9 节第 8 题中所引入的类似变量 $N_n(x)$, 亦可参阅本章第 9 节第 3 题. 两个变量 $L_n(x)$ 与 $N_n(x)$ 通常被称为时间集合 $\{k:\ 0 \leqslant k \leqslant n\}$ 中位于状态 x 的局部时.)

证明, 对于 $k = 0, 1, \cdots, n$, 有

$$\mathbf{P}\{L_{2n}(0) = k\} = \mathbf{P}\{L_{2n+1}(0) = k\} = 2^{-2n+k} C_{2n-k}^{n},$$
$$\mathbf{P}\{L_n(0) = k\} = 2^{-2[n/2]+k} C_{2[n/2]-k}^{[n/2]},$$
$$\mathbf{P}\{L_{2n}(0) < k\} = \mathbf{P}\{\sigma_k(0) > 2n\} = 2^{-2n} \sum_{j=0}^{k-1} 2^j C_{2n-j}^{n},$$
$$\mathbf{P}\{L_n(x) = 0\} = \mathbf{P}\{\sigma_1(x) > n\} = \sum_{j=n+1}^{\infty} 2^{-j} \frac{x}{j} C_j^{(j+x)/2},$$

而对于 $x = \pm 1, \pm 2, \cdots$, 有

$$\mathbf{P}\{L_{\sigma_1(0)}(x) = 0\} = \frac{2|x| - 1}{2|x|}, \quad \mathbf{E} L_{\sigma_1(0)}(x) = 1.$$

(变量 $\sigma_1(x)$ 的定义见第 9 题.)

12. 在上题中的条件下, 令

$$\mu(n) = \min\Big\{k,\ 0 \leqslant k \leqslant n:\ \ S_k = \max_{0\leqslant j\leqslant k} S_j\Big\}.$$

证明

$$\mathbf{P}\{\mu(2n) = k\} = \begin{cases} C_{2[k/2]}^{[k/2]} \cdot C_{2n-2[k/2]}^{n-[k/2]} \cdot 2^{-2n-1}, & k = 1, 2, 3, \cdots, 2n; \\ C_{2n}^{n} \cdot 2^{-2n}, & k = 0. \end{cases}$$

§11. 鞅. 鞅对随机游动的某些应用

1. 设 $\mathscr{D}_0 \preccurlyeq \mathscr{D}_1 \preccurlyeq \cdots \preccurlyeq \mathscr{D}_n$ 为分割序列, $\mathscr{D}_0 = \{\Omega\}$; η_k 为 $\mathscr{D}_k$ 可测的变量, $1 \leqslant k \leqslant n$. 证明, 按照如下方式定义的序列 $\xi = (\xi_k, \mathscr{D}_k)_{1\leqslant k\leqslant n}$ 是鞅:

$$\xi_k = \sum_{l=1}^{k} \Big(\eta_l - \mathbf{E}(\eta_l \mid \mathscr{D}_{l-1})\Big).$$

提示: 直接验证 $\mathbf{E}(\xi_{k+1} - \xi_k \mid \mathscr{D}_k) = 0$.

2. 设 $\eta_1, \cdots, \eta_n$ 为随机变量, 有 $\mathbf{E}\eta_1 = 0$, $\mathbf{E}(\eta_k \mid \eta_1, \cdots, \eta_{k-1}) = 0$, $1 \leqslant k \leqslant n$. 令 $\xi_1 = \eta_1$,

$$\xi_{k+1} = \sum_{i=1}^{k} f_i(\eta_1, \cdots, \eta_i)\eta_{i+1}, \quad k < n,$$

其中 $f_i(\eta_1, \cdots, \eta_i)$ 为某些函数. 证明, 序列 $\xi = (\xi_k)_{1\leqslant k\leqslant n}$ 为鞅.

3. 证明, 任何鞅序列 $\xi = (\xi_k, \mathscr{D}_k)_{1\leqslant k\leqslant n}$ 都具有增量不相关性, 即若 $a < b < c < d$, 那么就有

$$\mathbf{cov}(\xi_d - \xi_c, \xi_b - \xi_a) = 0.$$

(再次强调一下, 本章括号内的随机变量都只取有限个值.)

4. 设 $\xi = (\xi_1, \cdots, \xi_n)$ 为某个随机变量序列, 其中 ξ_k 为 $\mathscr{D}_k$ 可测 ($\mathscr{D}_1 \preccurlyeq \mathscr{D}_2 \preccurlyeq \cdots \preccurlyeq \mathscr{D}_n$). 证明, 为了该序列关于 $(\mathscr{D}_k)$ 是鞅, 必须且只需, 对于任何关于 $\mathscr{D}_k$ 的停时 τ, 都有 $\mathbf{E}\xi_\tau = \mathbf{E}\xi_1$. ("对于任何停时" 可以换为 "对于任何只取两个值的停时".)

提示: 设对于任何只取两个值的停时 τ, 都有 $\mathbf{E}\xi_\tau = \mathbf{E}\xi_1$. 取定 $k \in \{1, \cdots, n-1\}$, $A \in \mathscr{D}_k$, 考察停时

$$\tau(\omega) = \begin{cases} k, & \text{如果 } \xi_k(\omega) \notin A, \\ k+1, & \text{如果 } \xi_k(\omega) \in A. \end{cases}$$

证明 $\mathbf{E}\xi_\tau = \mathbf{E}\xi_k I_{\overline{A}} + \mathbf{E}\xi_k I_A$ 蕴涵 $\mathbf{E}\xi_{k+1} I_A = \mathbf{E}\xi_k I_A$.

5. 证明, 如果 $\xi = (\xi_k, \mathscr{D}_k)_{1\leqslant k\leqslant n}$ 是鞅, τ 是停时, 则对任何 $k \leqslant n$, 都有

$$\mathbf{E}[\xi_n I_{\{\tau=k\}}] = \mathbf{E}[\xi_k I_{\{\tau=k\}}].$$

6. 设 $\xi = (\xi_k, \mathscr{D}_k)_{1\leqslant k\leqslant n}$ 与 $\eta = (\eta_k, \mathscr{D}_k)_{1\leqslant k\leqslant n}$ 是两个鞅, 其中 $\xi_1 = \eta_1 = 0$. 证明

$$\mathbf{E}\xi_n\eta_n = \sum_{k=2}^{n} \mathbf{E}(\xi_k - \xi_{k-1})(\eta_k - \eta_{k-1}),$$

特别地, 有

$$\mathbf{E}\xi_n^2 = \sum_{k=2}^{n} \mathbf{E}(\xi_k - \xi_{k-1})^2.$$

7. 设 $\eta_1, \cdots, \eta_n$ 为独立同分布随机变量序列, 有 $\mathbf{E}\eta_i = 0$. 证明, 按照如下两种方式所定义的序列 $\xi = (\xi_k)_{1\leqslant k\leqslant n}$ 都是鞅:

$$\xi_k = \Big(\sum_{i=1}^{k} \eta_i\Big)^2 - k\mathbf{E}\eta_1^2 \quad \text{和} \quad \xi_k = \frac{\exp\{\lambda(\eta_1 + \cdots + \eta_k)\}}{(\mathbf{E}\exp\{\lambda\eta_1\})^k}.$$

8. 设 $\eta_1, \cdots, \eta_n$ 为独立同分布随机变量序列, 取值于 (有限) 集合 Y. 令 $f_0(y) = \mathbf{P}\{\eta_1 = y\} > 0$, $y \in Y$, 而 $f_1(y)$ 为非负函数, 有 $\sum\limits_{y \in Y} f_1(y) = 1$. 证明, 如下所定义的序列 $\xi = (\xi_k, \mathscr{D}_k^\eta)_{1 \leqslant k \leqslant n}$ 是鞅:

$$\xi_k = \frac{f_1(\eta_1) \cdots f_1(\eta_k)}{f_0(\eta_1) \cdots f_0(\eta_k)}, \quad \mathscr{D}_k^\eta = \mathscr{D}_{\eta_1, \cdots, \eta_k}.$$

(变量 ξ_k 称为似然比, 在数理统计中起着重要作用.)

9. 将序列 $\xi = (\xi_k, \mathscr{D}_k)_{0 \leqslant k \leqslant n}$ 称为上鞅 (下鞅), 如果 $\mathbf{P}$-a.s. 地有

$$\mathbf{E}(\xi_{k+1} \mid \mathscr{D}_k) \leqslant \xi_k \quad (\geqslant \xi_k), \quad 0 \leqslant k \leqslant n.$$

证明, 任何上鞅 (下鞅) 都可以 (唯一地) 表示为

$$\xi_k = m_k - a_k \quad (+a_k)$$

的形式, 其中 $m = (m_k, \mathscr{D}_k)_{0 \leqslant k \leqslant n}$ 是鞅, 而 $a = (a_k, \mathscr{D}_k)_{0 \leqslant k \leqslant n}$ 是这样的非降序列, 其中 $a_0 = 0$, 而对于所有 $k \geqslant 1$, 变量 a_k 都是 $\mathscr{D}_{k-1}$ 可测的.

10. 设 $\xi = (\xi_k, \mathscr{D}_k)_{0 \leqslant k \leqslant n}$ 与 $\eta = (\eta_k, \mathscr{D}_k)_{0 \leqslant k \leqslant n}$ 是上鞅, 而 τ 是关于分割 $(\mathscr{D}_k)_{0 \leqslant k \leqslant n}$ 的某个停时, 并且有 $\mathbf{P}\{\xi_\tau \geqslant \eta_\tau\} = 1$. 证明, 按照如下两种方式 "转换" 而成的序列 $\zeta = (\zeta_k, \mathscr{D}_k)_{0 \leqslant k \leqslant n}$ 都是上鞅:

$$\zeta_k = \xi_k I(\tau > k) + \eta_k I(\tau \leqslant k)$$

或

$$\zeta_k = \xi_k I(\tau \geqslant k) + \eta_k I(\tau < k).$$

11. 设 $\xi = (\xi_k, \mathscr{D}_k)_{0 \leqslant k \leqslant n}$ 是下鞅, 有

$$\xi_k = \sum_{m \leqslant k} I_{A_m},$$

其中 $A_m \in \mathscr{D}_m$. 试求该下鞅的杜布分解.

12. 设 $\xi = (\xi_k, \mathscr{D}_k)_{1 \leqslant k \leqslant n}$ 是下鞅[①], 证明如下的最大不等式:

$$\mathbf{E} \max_{l \leqslant n} \xi_l^+ \leqslant \frac{e}{1-e}\Big(1 + \mathbf{E}(\xi_n^+ \ln^+ \xi_n^+)\Big),$$

其中 $\ln^+ x = \max(\ln x, 0)$.

① 原文为: 设 $\xi = (\xi_k, \mathscr{D}_k)_{1 \leqslant k \leqslant n}$ —— 译者注.

§12. 马尔可夫链. 遍历性定理. 强马尔可夫性

1. 设 $\xi = (\xi_0, \xi_1, \cdots, \xi_n)$ 为取值于 X 的马尔可夫链, $f = f(x)$ $(x \in X)$ 为某个函数. 试问, $(f(\xi_0), f(\xi_1), \cdots, f(\xi_n))$ 可否形成马尔可夫链? "反向" 序列 $(\xi_n, \xi_{n-1}, \cdots, \xi_0)$ 可否为马尔可夫链?

2. 设 $\mathbb{P} = \|p_{ij}\|$, $1 \leqslant i, j \leqslant r$, 为随机矩阵, λ 是它的一个特征值, 亦即方程 $\det\|\mathbb{P} - \lambda E\| = 0$ 的根. 证明, $\lambda_1 = 1$ 是一个特征值, 而其余的根 $\lambda_2, \cdots, \lambda_r$ 的绝对值都不超过 1. 在此, 如果存在 n, 使得 $\mathbb{P}^n > 0$ (意即所有 $p_{ij}^{(n)} > 0$), 则有 $|\lambda_i| < 1$, $i = 2, \cdots, r$. 再证明, 如果所有 $\lambda_1, \cdots, \lambda_r$ 互不相等, 则转移概率 $p_{ij}^{(k)}$ 满足关系式

$$p_{ij}^{(k)} = \pi_j + a_{ij}(2)\lambda_2^k + \cdots + a_{ij}(r)\lambda_r^k,$$

其中 $\pi_j, a_{ij}(2), \cdots, a_{ij}(r)$ 均可通过矩阵 $\mathbb{P}$ 的元素来表示. (通过这样一个代数变换, 可以得到关于马尔可夫链的渐近性质的分析, 特别地, 可以推出, 如果 $|\lambda_2| < 1, \cdots, |\lambda_r| < 1$, 则对每个 j, 都存在不依赖于 i 的极限 $\lim\limits_k p_{ij}^{(k)}$.)

3. 设 $\xi = (\xi_0, \xi_1, \cdots, \xi_n)$ 为具有 (有限) 状态空间 X 的齐次马尔可夫链, 其转移概率矩阵为 $\mathbb{P} = \|p_{xy}\|$. 我们记

$$T_\varphi(x) = \mathbf{E}[\varphi(\xi_1) \mid \xi_0 = x] \quad \left(= \sum_y \varphi(y) p_{xy}\right)$$

为一步转移算子. 设非负函数 $\varphi = \varphi(x)$ 满足方程

$$T_\varphi(x) = \varphi(x), \quad x \in X,$$

亦即为 "调和的". 证明, 随机变量序列

$$\zeta = (\zeta_k, \mathscr{D}_k^\xi)_{0 \leqslant k \leqslant n}$$

形成鞅, 其中 $\zeta_k = \varphi(\xi_k)$, $\mathscr{D}_k^\xi = \mathscr{D}_{\xi_0, \cdots, \xi_k}$.

4. 设 $\xi = (\xi_n, \Pi, \mathbb{P})$ 和 $\widetilde{\xi} = (\widetilde{\xi}_n, \widetilde{\Pi}, \mathbb{P})$ 是两个马尔可夫链, 具有不同的起始分布 $\Pi = (p_1, \cdots, p_r)$ 和 $\widetilde{\Pi} = (\tilde{p}_1, \cdots, \tilde{p}_r)$. 记 $\Pi^{(n)} = (p_1^{(n)}, \cdots, p_r^{(n)})$ 和 $\widetilde{\Pi}^{(n)} = (\tilde{p}_1^{(n)}, \cdots, \tilde{p}_r^{(n)})$. 证明, 如果 $\min\limits_{i,j} p_{ij} \geqslant \varepsilon > 0$, 则

$$\sum_{i=1}^r |\tilde{p}_i^{(n)} - p_i^{(n)}| \leqslant 2(1 - r\varepsilon)^n.$$

提示: 采用数学归纳法, 对 n 归纳.

5. 设 P 与 Q 为两个随机矩阵. 证明, PQ 和 $\alpha P + (1-\alpha)Q$ 也是随机矩阵, 其中 $0 \leqslant \alpha \leqslant 1$.

6. 我们来考察齐次马尔可夫链 $(\xi_0, \xi_1, \cdots, \xi_n)$, 其状态空间为 $X = \{0, 1\}$, 转移概率矩阵为

$$\begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix},$$

其中 $0 < \alpha < 1$, $0 < \beta < 1$. 令 $S_n = \xi_0 + \xi_1 + \cdots + \xi_n$. 推广棣莫弗–拉普拉斯定理 (第 6 节), 证明

$$\mathbf{P}\left\{\frac{S_n - \frac{\alpha}{\alpha+\beta}n}{\sqrt{\frac{n\alpha\beta(2-\alpha-\beta)}{(\alpha+\beta)^3}}} \leqslant x\right\} \to \Phi(x), \quad n \to \infty.$$

并验证: 如果 $\alpha + \beta = 1$, 而且 $\xi_0, \xi_1, \cdots, \xi_n$ 相互独立, 则上述表达式演变为

$$\mathbf{P}\left\{\frac{S_n - \alpha n}{\sqrt{n\alpha\beta}} \leqslant x\right\} \to \Phi(x), \quad n \to \infty.$$

7. 设 $\xi_0, \xi_1, \cdots, \xi_N$ 为相互独立的伯努利随机变量序列, 有 $\mathbf{P}\{\xi_i = 1\} = \mathbf{P}\{\xi_i = -1\} = 1/2$. 定义 $\eta_0 = \xi_0$, $\eta_n = \frac{\xi_{n-1}+\xi_n}{2}$, $1 \leqslant n \leqslant N$. 试问:

(a) 序列 $(\eta_0, \eta_1, \cdots, \eta_N)$ 是否为马尔可夫链?

(b) 如果令 $\zeta_0 = \xi_0$, $\zeta_n = \xi_{n-1}\xi_n$, $1 \leqslant n \leqslant N$, 那么序列 $(\zeta_0, \zeta_1, \cdots, \zeta_N)$ 是否为马尔可夫链?

8. 设 $X_1, \cdots, X_n$ 为独立同分布随机变量, 将它们排为非降变列 (即按非降顺序排列), 得到顺序 (秩序) 统计量 $X_1^{(n)}, \cdots, X_n^{(n)}$. (即 $X_1^{(n)} = \min\{X_1, \cdots, X_n\}, \cdots, X_n^{(n)} = \max\{X_1, \cdots, X_n\}$. 如果有 $X_{i_1} = \cdots = X_{i_k} = \min\{X_1, \cdots, X_n\}$ [①], 而 $i_1 < \cdots < i_k$, 则以 X_{i_1} 作为 $X_1^{(n)}$. 这种约定亦在其他场合下被采用, 可参阅第二章第 8 节第 19 题.)

尽管 $X_1, \cdots, X_n$ 相互独立, 它们的顺序统计量 $X_1^{(n)}, \cdots, X_n^{(n)}$ 却是相依的.

证明, 如果 X_i 一共只取两个不同的可能值, 则它们的顺序统计量形成马尔可夫链. 并举例说明, 当 X_i 可取三个不同值时, 这一结论未必成立. (我们指出, 如果 X_i 的分布连续, 则变差序列恒为马尔可夫链, 参阅第二章第 3 节.)

① 原文为: 如果有 $X_{i_1} = \cdots = X_{i_k} = \min\{X_1, \cdots, X_k\}$ —— 译者注.

第二章　概率论的数学基础

§1. 有无限种结局试验的概率模型. 柯尔莫戈洛夫公理化体系

1. 设 $\Omega = \{r: r \in [0,1]\}$ 为 $[0,1]$ 上的有理数集, $\mathscr{A}$ 为代数, 其中的每个集合都是有限个互不相交的如下形式的集合 A 的和: $\{r: a < r < b\}$, $\{r: a \leqslant r < b\}$, $\{r: a < r \leqslant b\}$, $\{r: a \leqslant r \leqslant b\}$, 而 $\mathbf{P}(A) = b - a$. 证明, $\mathbf{P}(A)$, $A \in \mathscr{A}$ 是集合的有限可加函数, 但不是可数可加的.

2. 设 Ω 为某个可数集, $\mathscr{F}$ 为它的子集的全体. 令 $\mu(A) = 0$, 如果 A 为有限子集; $\mu(A) = \infty$, 如果 A 为无限子集. 证明, μ 为有限可加的, 但不是可数可加的.

3. 设 μ 为 $(\Omega, \mathscr{F})$ 上的可数可加的测度, 证明,

(a) 如果 $A_n \uparrow A$, 则 $\mu(A_n) \uparrow \mu(A)$.

(b) 如果 $A_n \downarrow A$, 并且对某个 k, 有 $\mu(A_k) < \infty$, 则 $\mu(A_n) \downarrow \mu(A)$.

(c) 如果 μ 为有限测度 $(\mu(\Omega) < \infty)$, 而 $A = \lim A_n$, 意即 $A = \overline{\lim} A_n = \underline{\lim} A_n$, 则 $\mu(A) = \lim \mu(A_n)$.

(进一步的讨论见第 15 题.)

提示: 利用表达式

$$\underline{\lim} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k, \quad \overline{\lim} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k.$$

4. 验证对称差的下列性质:

$$(A \triangle B) \triangle C = A \triangle (B \triangle C),$$
$$(A \triangle B) \triangle (B \triangle C) = A \triangle C,$$
$$A \triangle B = C \Leftrightarrow A = B \triangle C.$$

5. 证明, 按照如下方式定义的 "距离" $\rho_1(A,B)$ 和 $\rho_2(A,B)$ 满足三角形不等式:

$$\rho_1(A,B) = \mathbf{P}(A \triangle B),$$
$$\rho_2(A,B) = \begin{cases} \dfrac{\mathbf{P}(A \triangle B)}{\mathbf{P}(A \cup B)}, & \text{如果 } \mathbf{P}(A \cup B) \neq 0, \\ 0, & \text{如果 } \mathbf{P}(A \cup B) = 0, \end{cases}$$

其中 $A \triangle B$ 是集合 A 与 B 的对称差.

提示: 利用包含关系 $A \triangle C \subseteq (A \triangle B) \cup (B \triangle C)$.

6. 设 μ 为代数 $\mathscr{A}$ 上的有限可加测度, 集合 $A_1, A_2, \cdots \in \mathscr{A}$ 两两不交, 且 $A = \sum\limits_{i=1}^{\infty} A_i \in \mathscr{A}$. 证明, $\mu(A) \geqslant \sum\limits_{i=1}^{\infty} \mu(A_i)$.

7. 证明,

$$\overline{\limsup A_n} = \liminf \overline{A}_n, \quad \overline{\liminf A_n} = \limsup \overline{A}_n,$$
$$\liminf A_n \subseteq \limsup A_n, \quad \limsup (A_n \cup B_n) = \limsup A_n \cup \limsup B_n,$$
$$\liminf (A_n \cap B_n) = \liminf A_n \cap \liminf B_n,$$
$$\limsup A_n \cap \liminf B_n \subseteq \limsup (A_n \cap B_n) \subseteq \limsup A_n \cap \limsup B_n.$$

如果 $A_n \uparrow A$ 或 $A_n \downarrow A$, 则

$$\liminf A_n = \limsup A_n.$$

8. 设 (x_n) 为实数列, $A_n = (-\infty, x_n)$. 证明, 在 $x = \limsup x_n$ 与 $A = \limsup A_n$ 之间存在关系 $(-\infty, x) \subseteq A \subseteq (-\infty, x]$. (换言之, A 等于 $(-\infty, x)$ 或 $(-\infty, x]$.)

9. 设 $A_1, A_2, \cdots$ 为集合 Ω 的子集序列. 证明,

$$\limsup (A_n \backslash A_{n+1}) = \limsup (A_{n+1} \backslash A_n) = (\limsup A_n) \setminus (\liminf A_n).$$

10. 试举例说明, 对于可以取值 $+\infty$ 的测度, 一般来说, 不能由它的可数可加性推出它在空集 $\varnothing$ 处的连续性.

11. 设 $A_1, \cdots, A_n$ 为 $\mathscr{F}$ 中的某些事件. 称该事件组为*可交换的*(exchangeable 或 interchangeable), 如果对任何 $1 \leqslant l \leqslant n$ 和任何 $1 \leqslant i_1 < \cdots < i_l \leqslant n$, 都有 $\mathbf{P}(A_{i_1} \cap \cdots \cap A_{i_l})$ 彼此相同 $(= p_l)$. 证明, 对于这样的事件组, 有如下的 (容斥) 公式成立 (试与第一章第 1 节第 12 题作比较):

$$\mathbf{P}\left(\bigcup_{i=1}^{n} A_i\right) = np_1 - C_n^2 p_2 + C_n^3 p_3 - \cdots + (-1)^n p_n.$$

12. 设 $(A_k)_{k\geqslant 1}$ 为可交换事件的无穷序列, 亦即对任何 $n \geqslant 1$, 任何数组 $1 \leqslant i_1 < \cdots < i_l \leqslant n$, 都有 $\mathbf{P}(A_{i_1} \cap \cdots \cap A_{i_l})$ 彼此相同 $(= p_l)$. 证明,

$$\mathbf{P}\left(\varliminf_n A_n\right) = \mathbf{P}\left(\bigcap_{i=1}^{\infty} A_i\right) = \lim_{l \to \infty} p_l,$$
$$\mathbf{P}\left(\varlimsup_n A_n\right) = \mathbf{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = 1 - \lim_{l \to \infty} (-1)^l \Delta^l(p_0),$$

其中 $p_0 = 1$, $\Delta^1(p_n) = p_{n+1} - p_n$, $\Delta^l(p_n) = \Delta^1(\Delta^{l-1}(p_n))$, $l \geqslant 2$.

13. 设 $(A_n)_{n\geqslant 1}$ 为某个集合序列, $I(A)$ 是集合 A 的示性函数. 证明,

(a) $I\left(\varliminf_n A_n\right) = \varliminf_n I(A_n) \quad I\left(\varlimsup_n A_n\right) = \varlimsup_n I(A_n)$.

(b) $\varlimsup_n I(A_n) - \varliminf_n I(A_n) = I\left(\varlimsup_n A_n \setminus \varliminf_n A_n\right)$.

(c) $I\left(\bigcup\limits_{n=1}^{\infty} A_n\right) \leqslant \sum\limits_{n=1}^{\infty} I(A_n)$.

14. 证明

$$I\left(\bigcup_{n=1}^{\infty} A_n\right) = \max_{n\geqslant 1} I(A_n), \quad I\left(\bigcap_{n=1}^{\infty} A_n\right) = \min_{n\geqslant 1} I(A_n).$$

15. (续第 3 题) 设 μ 为 $(\Omega, \mathscr{F})$ 上的可数可加的测度, 证明,

(a) $\mu(\varliminf A_n) \leqslant \varliminf \mu(A_n)$.

(b) 如果 μ 为有限测度 $(\mu(\Omega) < \infty)$, 则

$$\mu(\varlimsup A_n) \geqslant \varlimsup \mu(A_n).$$

(c) 对于概率测度 $(\mu = \mathbf{P})$, 有 (*关于集合的法图引理*):

$$\mathbf{P}(\varliminf A_n) \leqslant \varliminf \mathbf{P}(A_n) \leqslant \varlimsup \mathbf{P}(A_n) \leqslant \mathbf{P}(\varlimsup A_n).$$

由此推出关于概率 $\mathbf{P}$ 的 "连续" 性质 2) 与 3) (第 2 节中的定理) 的推广: 如果 $A = \lim\limits_n A_n$(即 $\varlimsup A_n = \varliminf A_n = A$), 则 $\mathbf{P}(A) = \lim\limits_n \mathbf{P}(A_n)$.

16. 记 $A^* = \varlimsup A_n$, $A_* = \varliminf A_n$. 证明, $\mathbf{P}(A_n - A_*) \to 0$, $\mathbf{P}(A^* - A_n) \to 0$.

17. 设集合 $A_n \to A$ (意即 $A = A^* = A_*$, 参阅第 16 题). 证明, $\mathbf{P}(A \triangle A_n) \to 0$.

18. 设集合 A_n 按如下意义收敛到 A : $\mathbf{P}(A \triangle \overline{\lim} A_n) = \mathbf{P}(A \triangle \underline{\lim} A_n) = 0$. 证明, $\mathbf{P}(A \triangle A_n) \to 0$.

19. 设 $A_0, A_1, \cdots$ 与 $B_0, B_1, \cdots$ 都是集合 Ω 的子集. 证明, 对称差满足下列性质:

$$A_0 \triangle B_0 = \overline{A}_0 \triangle \overline{B}_0,$$

$$A_0 \triangle \left(\bigcup_{n\geqslant 1} B_n\right) \subseteq \bigcup_{n\geqslant 1} (A_0 \triangle B_n),$$

$$A_0 \triangle \left(\bigcap_{n\geqslant 1} B_n\right) \supseteq \bigcap_{n\geqslant 1} (A_0 \triangle B_n),$$

$$\left(\bigcup_{n\geqslant 1} A_n\right) \triangle \left(\bigcup_{n\geqslant 1} B_n\right) \subseteq \bigcup_{n\geqslant 1} (A_n \triangle B_n),$$

$$\left(\bigcap_{n\geqslant 1} A_n\right) \triangle \left(\bigcap_{n\geqslant 1} B_n\right) \subseteq \bigcup_{n\geqslant 1} (A_n \triangle B_n).$$

20. 设 A, B, C 为随机事件, 证明

$$|\mathbf{P}(A \cap B) - \mathbf{P}(A \cap C)| \leqslant \mathbf{P}(B \triangle C).$$

21. 证明, 对于任意三个事件 A, B 和 C, 它们之中恰好发生一个事件的概率等于

$$[\mathbf{P}(A) + \mathbf{P}(B) + \mathbf{P}(C)] = 2[\mathbf{P}(AB) + \mathbf{P}(AC) + \mathbf{P}(BC)] + 3\mathbf{P}(ABC).$$

(试比较第一章第 1 节第 13 题.)

22. 设 $(A_n)_{n\geqslant 1}$ 是 $\mathscr{F}$ 中的某个事件序列, 使得

$$\sum_n \mathbf{P}(A_n \;\triangle\; A_{n+1}) < \infty.$$

证明,

$$\mathbf{P}\{(\overline{\lim} A_n) \;\triangle\; (\underline{\lim} A_n)\} = 0.$$

23. 证明, 等于任何两个事件 A 和 B, 都有

$$\max(\mathbf{P}(A),\ \mathbf{P}(B)) \leqslant \mathbf{P}(A \cup B) \leqslant 2\max(\mathbf{P}(A),\ \mathbf{P}(B))$$

和

$$\mathbf{P}(A \cup B)\mathbf{P}(A \cap B) \leqslant \mathbf{P}(A)\mathbf{P}(B).$$

何时等号成立?

24. 设 $(A_n)_{n\geqslant 1}$ 与 $(B_n)_{n\geqslant 1}$ 是两个事件序列, 并且对每个 $n \geqslant 1$, 都有 $A_n \subseteq B_n$, 证明 $\{A_n \text{ i.o.}\} \subseteq \{B_n \text{ i.o.}\}$.[1]

25. 仍设 $(A_n)_{n\geqslant 1}$ 与 $(B_n)_{n\geqslant 1}$ 是两个事件序列, 并且

$$\mathbf{P}\{A_n \text{ i.o.}\} = 1, \quad \mathbf{P}\{\overline{B}_n \text{ i.o.}\} = 0.$$

证明 $\mathbf{P}\{A_n \cap B_n \text{ i.o.}\} = 1$.

26. 试举出两个有限测度 μ_1 和 μ_2 的例子 (即有 $\mu_1(\Omega) < \infty$, $\mu_2(\Omega) < \infty$), 使得满足条件 $\nu \geqslant \mu_1$, $\nu \geqslant \mu_2$ 的最小测度 ν 不是 $\max(\mu_1,\ \mu_2)$, 而是 $\mu_1 + \mu_2$.

27. 设在可测空间 $(\Omega, \mathscr{F})$ 中给定了一列概率测度 $\mathbf{P}_1, \mathbf{P}_2, \cdots$, 使得对每个 $A \in \mathscr{F}$, 都有

$$\mathbf{P}_n(A) \to \mathbf{P}(A),$$

其中 $\mathbf{P}(A)$ 是 $A \in \mathscr{F}$ 上的某个集合函数. 试证明如下结果 (布塔尔–汗–萨克斯定理):

(a) $\mathbf{P} = \mathbf{P}(\cdot)$ 是 $(\Omega, \mathscr{F})$ 上的概率测度.

(b) 对于 $\mathscr{F}$ 中的任何满足条件 $A_k \downarrow \varnothing\,(k \to \infty)$ 的集合序列 $A_1, A_2, \cdots$, 都有 $\sup\limits_n \mathbf{P}_n(A_k) \downarrow 0,\ \ k \to \infty$.

28. 试在 $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ 上构造一列测度 $\mu_n = \mu_n(A)$, $A \in \mathscr{B}(\mathbb{R})$, $n \geqslant 1$, 使得对于每个 $A \in \mathscr{B}(\mathbb{R})$, 序列 $(\mu_n(A))_{n\geqslant 1}$ 都是下降的, 然而 $\nu(A) = \lim\limits_n \mu_n(A)$, $A \in \mathscr{B}(\mathbb{R})$, 却不具有可数可加性, 从而不是测度.

29. 设 $(\Omega, \mathscr{F}, \mathbf{P})$ 为概率空间, $(A_n)_{n\geqslant 1}$ 为一列事件. 设 $\mathbf{P}(A_n) \geqslant c > 0, n \geqslant 1$, 及 $A = \overline{\lim} A_n$. 证明 $\mathbf{P}(A) \geqslant c$.

30. (惠更斯问题)A, B 二人轮流抛掷一对骰子. 如果 A 在 B 抛出 7 点之前抛出 6 点, 则 A 赢, 否则就是 B 赢. 如果 A 先开始, 试求他赢的概率 P_A.

提示: 以 A_k 表示 A 在第 $k+1$ 次抛掷时取胜的事件, 则有 $P_A = \sum\limits_{k=0}^{\infty} \mathbf{P}(A_k)$.

(本题答案为: $P_A = 30/61$.)

31. 设 Ω 为由至多可数个元素构成的集合, $\mathscr{F}$ 为它的子集形成的某个 σ-代数. 证明, 一定存在一个分割 $\mathscr{D} = \{D_1, D_2, \cdots\}$ ($\bigcup\limits_{i\in N} D_i = \Omega$, $D_i \cap D_j = \varnothing$, $i \neq j$ 和 $N = \{1, 2, \cdots\}$), 使得 $\mathscr{F}$ 可由 $\mathscr{D}$ 生成:

$$\mathscr{F} = \left\{\bigcup_{i\in M} D_i,\ \ M \subseteq N\right\}.$$

(试与第一章第 1 节第 3 小节关于有限集合 Ω 情形下的断言作比较.)

[1] i.o.表示无限多次 (infinitely often) —— 译者注.

提示: 可以通过在 Ω 中考察由如下的等价关系 "$\omega_1 \sim \omega_2$" 所形成的等价类来构造 $\mathscr{D}$:

$$\omega_1 \sim \omega_2 \quad \Longleftrightarrow \quad \{\omega_1 \in A \Leftrightarrow \omega_2 \in A \text{ 对于每个 } A \in \mathscr{F}\}.$$

§2. 代数和 σ-代数. 可测空间

1. 设 $\mathscr{B}_1$ 与 $\mathscr{B}_2$ 是空间 Ω 的子集的两个 σ-代数. 试问: 如下两个集合类是否也是 σ-代数:

$$\mathscr{B}_1 \cap \mathscr{B}_2 \equiv \{A:\ A \in \mathscr{B}_1 \text{ 且 } A \in \mathscr{B}_2\},$$
$$\mathscr{B}_1 \cup \mathscr{B}_2 \equiv \{A:\ A \in \mathscr{B}_1 \text{ 或 } A \in \mathscr{B}_2\}?$$

设 $\mathscr{B}_1 \vee \mathscr{B}_2$ 是由 σ-代数 $\mathscr{B}_1$ 和 $\mathscr{B}_2$ 中的集合所张成的最小 σ-代数 $\sigma(\mathscr{B}_1, \mathscr{B}_2)$. 证明, $\mathscr{B}_1 \vee \mathscr{B}_2$ 重合于由所有形如 $B_1 \cap B_2$ 的集合所生成的 σ-代数, 其中 $B_1 \in \mathscr{B}_1$ 而 $B_2 \in \mathscr{B}_2$.

提示: 验证 $\mathscr{B}_1 \cap \mathscr{B}_2$ 是 σ-代数, 举例说明 $\mathscr{B}_1 \cup \mathscr{B}_2$ 不是 σ-代数. (设空间 Ω 为一个 3 点集, 即可构造出所需的例子.)

2. 设 $\mathscr{D} = \{D_1, D_2, \cdots\}$ 是 Ω 的某个可数分割, 而 $\mathscr{B} = \sigma(\mathscr{D})$. 试问 σ-代数 $\mathscr{B}$ 的势为多少?

提示: 可以确认 σ-代数 $\mathscr{B}$ 具有连续统的势, 因为可以把集合 $D^x = D_1^{x_1} \cup D_2^{x_2} \cup \cdots$, 其中 $x_i = 0$ 或 1 (若 $x_i = 0$, 则 $D_i^{x_i} = \varnothing$, 而若 $x_i = 1$, 则 $D_i^{x_i} = D_i$) 视为由 0 和 1 所构成的序列 $x = (x_1, x_2, \cdots)$, 这类序列的全体具有连续统的势.

3. 证明

$$\mathscr{B}(\mathbb{R}^n) \otimes \mathscr{B}(\mathbb{R}) = \mathscr{B}(\mathbb{R}^{n+1}).$$

4. 证明, 集合 (b) ~ (f) (参阅第 4 小节) 属于 $\mathscr{B}(\mathbb{R}^\infty)$.

提示: 为了证明, 例如, 性质 (b), 应当指出

$$\left\{x:\ \overline{\lim_n}\, x_n \leqslant a\right\} = \bigcap_{k=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \left\{x:\ x_n < a + \frac{1}{k}\right\}.$$

(换言之, $\overline{\lim_n}\, x_n \leqslant a \ \Leftrightarrow\ \forall\, k \in N,\ \exists\, m \in N:$ 当 $n \geqslant m$ 时, 有 $x_n < a + \frac{1}{k}$.) 类似可证性质 (c) ~ (f).

5. 证明, 集合 A_2 与 A_3 (参阅第 5 小节) 不属于 $\mathscr{B}(\mathbb{R}^\infty)$.

提示: 与证明集合 A_1 的情形类似, 关于 A_2 与 A_3 的证明也采用反证法.

6. 证明, 函数 (18) 事实上给出了一个测度.

7. 证明, $\mathscr{B}_0(\mathbb{R}^n) = \mathscr{B}(\mathbb{R}^n)$, $n \geqslant 1$, 并且 $\mathscr{B}_0(\mathbb{R}^\infty) = \mathscr{B}(\mathbb{R}^\infty)$.

8. 设 $\mathcal{C} = \mathcal{C}[0, \infty)$ 为定义在 $t \geqslant 0$ 上的所有连续函数 $x = x(t)$ 的空间. 证明, 关于距离

$$\rho(x,\ y) = \sum_{n=1}^{\infty} 2^{-n} \min\Big(\sup_{0 \leqslant t \leqslant n} |x_t - y_t|,\ 1\Big), \quad x, y \in \mathcal{C},$$

该空间 (与 $\mathcal{C} = \mathcal{C}[0,1]$ 一样) 是一个波利希空间, 亦即可分完备的距离空间, 并且由它的所有开集所生成的 σ-代数 $\mathscr{B}_0(\mathcal{C})$ 重合于由所有的柱集所生成的 σ-代数 $\mathscr{B}(\mathcal{C})$.

9. 证明, 条件组 $\{(\lambda_a), (\lambda_b), (\lambda_c)\}$ 与条件组 $\{(\lambda_a), (\lambda_b'), (\lambda_c')\}$ 相互等价 (参阅《概率》第一卷第二章第 2 节定义 2 与注 2).

10. 试由定理 1 推出定理 2.①

11. 证明, 定理 3 的证明过程中的系 $\mathscr{L}$ 是 λ-系.②

12. 称 σ-代数是 "可数生成的" 或者 "可分的", 如果它是某个可数的集合类所生成的.

(a) 证明, $\Omega = [0, 1]$ 的博雷尔子集的 σ-代数 $\mathscr{B}$ 是可数生成的.

(b) 作为例子, 证明, 两个 σ-代数 $\mathscr{F}_1$ 与 $\mathscr{F}_2$ 可能如此: $\mathscr{F}_1 \subset \mathscr{F}_2$, 而 $\mathscr{F}_2$ 为可数生成, 但 $\mathscr{F}_1$ 却不是可数生成的.

13. 证明, 为了 σ-代数 $\mathscr{G}$ 是可数生成的, 必须且只需, 对于某个随机变量 X, 有 $\mathscr{G} = \sigma(X)$ ($\sigma(X)$ 的定义见第 4 节第 4 小节).

14. 证明, $(X_1, X_2, \cdots)$ 为独立随机变量族, 如果对于每个 $n \geqslant 1$, 都有 $\sigma(X_n)$ 与 $\sigma(X_1, \cdots, X_{n-1})$ 独立.

15. 设 $(\Omega, \mathscr{F}, \mathbf{P})$ 为完备的概率空间 (参阅第 3 节第 1 小节和第 34 题), $\mathscr{G}$ 是 $\mathscr{F}$ 的子 σ-代数 ($\mathscr{G} \subseteq \mathscr{F}$). 而 $(\mathscr{E}_n)_{n \geqslant 1}$ 是 $\mathscr{F}$ 的非升的子 σ-代数序列 ($\mathscr{E}_1 \supseteq \mathscr{E}_2 \supseteq \cdots \mathscr{E}_n \subseteq \mathscr{F},\ n \geqslant 1$). 假设所有的子 σ-代数都已经被 $\mathscr{F}$ 中的 $\mathbf{P}$ 测度为 0 的集合完备化了的. 那么, 乍一看来, 似乎应当有如下的关系式 (至少精确到零测集)

---

①《概率》第一卷第二章第 2 节中的定理 1 为: 如果 $\mathscr{A}$ 是代数, 则有

$$\mu(\mathscr{A}) = \sigma(\mathscr{A}),$$

其中 $\mu(\mathscr{A})$ 是包含 $\mathscr{A}$ 的最小单调类. 此处要求利用这一定理, 证明如下结论:

(a) 任何 π-λ-系 $\mathscr{E}$ 都是 σ-代数.

(b) 如果 $\mathscr{E}$ 是集合的 π-系, 则有 $\lambda(\mathscr{E}) = d(\mathscr{E}) = \sigma(\mathscr{E})$.

(c) 如果 $\mathscr{E}$ 是集合的 π-系, 而 $\mathscr{L}$ 是 λ-系, 使得 $\mathscr{E} \subseteq \mathscr{L}$, 则有 $\sigma(\mathscr{E}) \subseteq \mathscr{L}$ —— 译者注.

②定理 3 为: 设 $\mathscr{E}$ 是 $\mathscr{F}$ 中的集合所成的 π-系, 而 $\mathscr{H}$ 是这样的一些 $\mathscr{F}$-可测的实值函数所成的集合, 有: (h1) 如果 $A \in \mathscr{E}$, 则 $I_A \in \mathscr{H}$; (h2) 如果 $f \in \mathscr{H}$, $h \in \mathscr{H}$, 则 $f + h \in \mathscr{H}$; (h3) 如果 $h_n \in \mathscr{H}$, $n \geqslant 1$, $h_n \uparrow h$, 则 $h \in \mathscr{H}$. 那么在 $\mathscr{H}$ 中包含了所有 $\sigma(\mathscr{E})$-可测的有界实值函数. 在该定理证明中构造了集合类 $\mathscr{L} = \{A \in \mathscr{F}:\ I_A \in \mathscr{H}\}$, 此处要求证明 $\mathscr{L}$ 是 λ-系 —— 译者注.

成立:

$$\bigcap_n \sigma(\mathscr{G},\ \mathscr{E}_n) = \sigma\left(\mathscr{G},\ \bigcap_n \mathscr{E}_n\right), \tag{*}$$

或者表示为

$$\bigcap_n (\mathscr{G}\ \vee\ \mathscr{E}_n) = \mathscr{G}\ \vee\ \bigcap_n \mathscr{E}_n, \tag{**}$$

其中 $\mathscr{G}\ \vee\ \mathscr{E}_n \equiv \sigma(\mathscr{G},\ \mathscr{E}_n)$ 是由 $\mathscr{G}$ 和 $\mathscr{E}_n$ 中的集合所生成的最小 σ-代数; 任何两个完备化了的属于 $\mathscr{F}$ 中的 σ-代数 $\mathscr{H}_1$ 与 $\mathscr{H}_2$ (精确到零测集) 的相等, 是指: 对于每个 $A \in \mathscr{H}_1$, 都能找到某个 $B \in \mathscr{H}_2$ (反之, 对于每个 $B \in \mathscr{H}_2$, 都能找到某个 $A \in \mathscr{H}_1$), 使得 $\mathbf{P}(A \triangle B) = 0$.

然而, 下面的来自 [22] 的例子却表明, 事实并非如此, 一般来说, 取上运算 $(\vee)$ 与取交运算 $(\cap)$ 是不能交换的.

(a) 设 $\xi_0, \xi_1, \xi_2, \cdots$ 为独立伯努利随机变量, 有 $\mathbf{P}\{\xi_i = 1\} = \mathbf{P}\{\xi_i = -1\} = 1/2$. 令 $X_n = \xi_0\xi_1\cdots\xi_n$, 以及

$$\mathscr{G} = \sigma(\xi_1, \xi_2, \cdots), \quad \mathscr{E}_n = \sigma(X_k,\ k > n).$$

证明, 在这里有

$$\bigcap_n \sigma(\mathscr{G},\ \mathscr{E}_n) \neq \sigma\left(\mathscr{G},\ \bigcap_n \mathscr{E}_n\right).$$

提示: 验证, 随机变量 ξ_0 是关于 $\bigcap_n \sigma(\mathscr{G},\ \mathscr{E}_n) (= \sigma(\mathscr{G},\ \xi_0))$ 可测的, 但同时却与 $\sigma(\mathscr{G}, \bigcap_n \mathscr{E}_n) (= \mathscr{G})$ 中的事件独立. (进一步的讨论见第七章第 4 节第 25 题.)

(b) 对于 σ-代数的取上运算 $(\vee)$ 与取交运算 $(\cap)$ 的不可交换性, 可以 (不依赖于 (a) 中的断言) 更为简单地直接推出 (参阅 [131]): 设 ξ_1 与 ξ_2 如 (a) 中的随机变量 (即相互独立的对称的伯努利随机变量), 而 $\mathscr{E}_1 = \sigma(\xi_1)$, $\mathscr{E}_2 = \sigma(\xi_2)$, $\mathscr{G} = \sigma(\xi_1\xi_2)$, 则

$$(\mathscr{G} \vee \mathscr{E}_1) \cap (\mathscr{G} \vee \mathscr{E}_2) \neq \mathscr{G} \vee (\mathscr{E}_1 \cap \mathscr{E}_2) \quad 且 \quad (\mathscr{G} \cap \mathscr{E}_1) \vee (\mathscr{G} \cap \mathscr{E}_2) \neq \mathscr{G} \cap (\mathscr{E}_1 \vee \mathscr{E}_2).$$

16. 设 $\mathscr{A}_1$ 与 $\mathscr{A}_2$ 是两个相互独立的集合类, 并且都是 π-系. 证明, $\sigma(\mathscr{A}_1)$ 与 $\sigma(\mathscr{A}_2)$ 也是相互独立的. 试举出两个相互独立的非 π-系的集合类的例子, 对于它们, $\sigma(\mathscr{A}_1)$ 与 $\sigma(\mathscr{A}_2)$ 不相互独立.

17. 设 $\mathscr{L}$ 是 π-系, 证明, $(A, B \in \mathscr{L},\ A \cap B = \varnothing) \Rightarrow (A \cup B \in \mathscr{L})$.

18. 设 $\mathscr{F}_1$ 与 $\mathscr{F}_2$ 是 Ω 中的子集的两个 σ-代数. 令

$$d(\mathscr{F}_1,\ \mathscr{F}_2) = 4 \sup_{A_1 \in \mathscr{F}_1,\ A_2 \in \mathscr{F}_2} |\mathbf{P}(A_1A_2) - \mathbf{P}(A_1)\mathbf{P}(A_2)|.$$

证明, 这个用来刻画 $\mathscr{F}_1$ 与 $\mathscr{F}_2$ 之间的相关程度的变量具有下列性质:

(a) $0 \leqslant d(\mathscr{F}_1,\ \mathscr{F}_2) \leqslant 1$.

(b) $d(\mathscr{F}_1,\ \mathscr{F}_2) = 0$ 当且仅当 $\mathscr{F}_1$ 与 $\mathscr{F}_2$ 相互独立.

(c) $d(\mathscr{F}_1,\ \mathscr{F}_2) = 1$ 当且仅当 $\mathscr{F}_1$ 与 $\mathscr{F}_2$ 的交中含有概率等于 1/2 的事件.

19. 利用定理 1 的证明方法, 证明, 包含集合类 $\mathscr{E}$ 的 λ-系 $\lambda(\mathscr{E})$ 和 π-系 $\pi(\mathscr{E})$ 的存在性与唯一性.

20. 设 $\mathscr{A}$ 是集合的某个代数, 具有性质: 任何不交的集合序列 $(A_n)_{n\geqslant 1}$, 只要 $A_n \in \mathscr{A}$, 就有 $\bigcup_{n=1}^{\infty} A_n \in \mathscr{A}$. 证明, $\mathscr{A}$ 是 σ-代数.

21. 设 $(\mathscr{F}_n)_{n\geqslant 1}$ 是上升的 σ-代数序列, 即 $\mathscr{F}_n \subseteq \mathscr{F}_{n+1},\ n \geqslant 1$. 证明, $\bigcup_{n=1}^{\infty} \mathscr{F}_n$ (一般来说, 仅仅只能) 是代数.

22. 设 $\mathscr{F}$ 是代数 (或 σ-代数), 集合 C 不属于 $\mathscr{F}$. 我们来考察由 $\mathscr{F}$ 中的集合和集合 C 所生成的最小代数 (相应的, σ-代数). 证明, 该代数 (相应的, σ-代数) 由形如 $(A \cap C) \cup (B \cap \overline{C})$ 的集合构成, 其中 $A, B \in \mathscr{F}$.

23. 设 $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty\} \cup \{\infty\}$ 是***增广数轴直线***. 博雷尔 σ-代数 $\mathscr{B}(\overline{\mathbb{R}})$ 可以定义为由所有形如 $[-\infty, x]\ (x \in \mathbb{R})$ 的集合所生成的 σ-代数, 其中 $[-\infty, x] = \{-\infty\} \cup (-\infty, x]$ (试与第 2 小节中的定义相比较). 证明, 该 σ-代数 $\mathscr{B}(\overline{\mathbb{R}})$ 亦重合于下列任何一个集合类所生成的 σ-代数:

(a) $[-\infty, x),\ x \in \mathbb{R}$, 或

(b) $(x, \infty],\ x \in \mathbb{R}$, 其中 $(x, \infty] = (x, -\infty) \cup \{\infty\}$, 或

(c) 所有有界区间, $\{-\infty\}$ 和 $\{\infty\}$.

24. 考察可测空间 $(\mathcal{C}, \mathscr{B}_0(\mathcal{C}))$, 其中 $\mathcal{C} = \mathcal{C}[0,1]$ 是全体连续函数 $x = (x_t),\ t \in [0,1]$ 的集合, 而 $\mathscr{B}_0(\mathcal{C})$ 是由关于距离 $\rho(x, y) = \sup_{0\leqslant t\leqslant 1} |x_t - y_t|$ 的开集所生成的博雷尔集合类.

(a) 证明, 空间 $(\mathcal{C}, \mathscr{B}_0(\mathcal{C}))$ 是完备的.

(b) 证明, 空间 $(\mathcal{C}, \mathscr{B}_0(\mathcal{C}))$ 是可分的.

提示: 为了证明此处的结论, 考察伯恩斯坦多项式是有益的, 参阅第一章第 5 节.

(c) 证明, 由可导函数构成的子空间 $\mathcal{C}^{\mathrm{dif}}[0,1] \subset \mathcal{C}[0,1]$ 不是 $\mathcal{C}[0,1]$ 中的博雷尔集.

25. 设 $\mathscr{A}_0$ 是 Ω 中的某个 (非空的) 子集类. 证明, 由该子集类生成的代数 $\alpha(\mathscr{A}_0)$ 可以按照如下方式构造出来.

令 $\mathscr{A}_1 = \mathscr{A}_0 \cup \{\Omega,\ \varnothing\}$, 而对于 $n \geqslant 1$, 令

$$\mathscr{A}_{n+1} = \{A \cup \overline{B}:\ A, B \in \mathscr{A}_n\}.$$

则有 $\mathscr{A}_0 \subseteq \mathscr{A}_1 \subseteq \cdots \subseteq \mathscr{A}_n \subseteq \cdots$ 和

$$\alpha(\mathscr{A}_0) = \bigcup_{n=1}^{\infty} \mathscr{A}_n.$$

26. 如果 $\mathscr{A}_0$ 是 Ω 中的某个 (非空的) 子集类, 则第 25 题已经给出如何构造出由 $\mathscr{A}_0$ 所生成的 (最小) 代数 $\alpha(\mathscr{A}_0)$ 的具体办法.

现在我们按照类似的步骤定义下列集合类:

$$\mathscr{A}_1 = \mathscr{A}_0 \cup \{\Omega, \varnothing\}, \qquad \mathscr{A}_2 = \mathscr{A}_1 \cup \left\{\bigcup_{n=1}^{\infty} B_n : B_1, B_2, \cdots \in \mathscr{A}_1\right\},$$
$$\mathscr{A}_3 = \mathscr{A}_2 \cup \{\overline{B} : B \in \mathscr{A}_2\}, \quad \mathscr{A}_4 = \mathscr{A}_3 \cup \left\{\bigcup_{n=1}^{\infty} B_n : B_1, B_2, \cdots \in \mathscr{A}_3\right\},$$
$$\mathscr{A}_5 = \mathscr{A}_4 \cup \{\overline{B} : B \in \mathscr{A}_4\}, \quad \mathscr{A}_6 = \mathscr{A}_5 \cup \left\{\bigcup_{n=1}^{\infty} B_n : B_1, B_2, \cdots \in \mathscr{A}_5\right\},$$
$$\cdots\cdots \qquad \cdots\cdots$$

可能会认为集合类 $\mathscr{A}_\infty = \bigcup_{n=1}^{\infty} \mathscr{A}_n$ 就是由集合类 $\mathscr{A}_0$ 所生成的 (最小) σ-代数. 然而却并非如此, 事实上, 一般来说, 虽然恒有 $\mathscr{A}_\infty \subseteq \sigma(\mathscr{A}_0)$, 但是却有 $\mathscr{A}_\infty \neq \sigma(\mathscr{A}_0)$, 这就是说, 由上述步骤未必就能得到 σ-代数 $\sigma(\mathscr{A}_0)$. 试举出例子, 说明存在 $\sigma(\mathscr{A}_0)$ (严格) 大于 $\mathscr{A}_\infty$ 的情况.

提示: 可以将数轴 $\mathbb{R}$ 上的所有区间的类取为 $\mathscr{A}_0$.

我们还需指出, 即使由 $\mathscr{A}_\infty$ 出发, 再按照上述步骤做下去, 一般来说, 也不能得到 $\sigma(\mathscr{A}_0)$. 为了最终得到 $\sigma(\mathscr{A}_0)$, 必须进行超限归纳 ($\aleph_1$"次"); 关于康托尔基数 (势) $\aleph_\alpha$, 连续统的势 $\mathfrak{c}$ 以及连续统假设的势 ($\mathfrak{c} = \aleph_1$) 参阅 [79] 第一卷, 第 235 页, 或第二卷, 第 1068 页.

27. 上一题中的构造和结论都表明, σ-代数的结构可以是非常复杂的. 勒贝格曾经认为 "$\mathbb{R}^2$ 中的任何博雷尔集 B 在坐标轴之一上的投影一定是 $\mathbb{R}^1$ 中的博雷尔集". 1916 年, 苏斯林给出了一个反例, 推翻了这一命题. 试尝试构造一个相应的反例.

提示: 苏斯林指出, 对于平面上的某些具体的开集序列 $A_1, A_2, \cdots$, 它们的交 $\bigcap A_n$ 在坐标轴之一上的投影不是博雷尔集.

28. (*施佩纳引理*) 设 $A = \{1, \cdots, n\}$ 是含有 n 个元素的集合, $\{A_1, \cdots, A_K\}$ 是 A 中的这样一些子集的族, 其中任何一者都不是任何别者的子集. 证明, 这种子集的数目 K 满足不等式 $K \leqslant C_n^{[n/2]}$.

29. 设 $\mathscr{E}$ 是 Ω 中的某个子集类, $\mathscr{F} = \sigma(\mathscr{E})$ 是 $\mathscr{E}$ 所生成的最小 σ-代数. 设 $A \in \mathscr{F}$. 试利用恰当集合原理, 证明, 可以从 $\mathscr{E}$ 中挑出一个可数子类 (记为 $\mathscr{C}$), 使得 $A \in \sigma(\mathscr{C})$.

30. 在距离空间 $(E, \mathscr{E}, \rho)$(以 ρ 为距离) 中, 博雷尔集合的 σ-代数 $\mathscr{E}$ 定义为由所有开集的类所生成的 σ-代数 (参阅第三章第 1 节第 3 小节). 证明, 对于某些距离空间, 由所有开 "球" 的类所生成的 σ-代数 $\mathscr{E}_0$ 严格地小于 $\mathscr{E}(\mathscr{E}_0 \subset \mathscr{E})$.

31. 证明, 不存在具有可数个元素的势为阿列夫零 $\aleph_0$ 的 σ-代数 (其势与全体自然数的势相等). 事实上, 在 σ-代数的结构中, 或者仅有有限个元素 (参阅第一章第 1 节第 10 题), 或者必具有不可数的势. 根据下一题, $\mathbb{R}^n$ 中的全体博雷尔集合的势为连续统的势 $\mathfrak{c} = 2^{\aleph_0}$(即对于自然数集的全体子集的势).

32. 设 $\mathfrak{c} = 2^{\aleph_0}$ 为连续统的势. 证明, 全体博雷尔集合的势为 $\mathfrak{c}$, 而全体勒贝格集合的 σ-代数的势为 $2^{\mathfrak{c}}$.

33. 设 B 为赋有勒贝格测度 λ 的数轴直线上的博雷尔集. 将如下的极限称为集合 B 的密度:

$$D(B) = \lim_{T\to\infty} \frac{\lambda\{B \cap [-T, T]\}}{2T} \qquad (*)$$

(假定该极限存在).

(a) 举例说明, 存在集合 B, 其密度 $D(B)$ 没有定义 ($(*)$ 式中的极限不存在).

(b) 证明, 对于互不相交的博雷尔集合 B_1 和 B_2, 有

$$D(B_1 + B_2) = D(B_1) + D(B_2).$$

(c) 举例说明, 存在互不相交的博雷尔集合序列 $B_1, B_2, \cdots$, 它们都有密度, 但是却 (意外地) 有

$$D\left(\bigcup_{i=1}^{\infty} B_i\right) \neq \sum_{i=1}^{\infty} D(B_i).$$

34. (σ-代数的完备化) 设 $(\Omega, \mathscr{F}, \mathbf{P})$ 为概率空间. 我们说该空间是**完备的** (或 $\mathbf{P}$ 完备的, 或关于测度 $\mathbf{P}$ 是完备的), 如果由 $B \in \mathscr{F}$ 并且 $\mathbf{P}(B) = 0$ 可以推出所有集合 $A \subseteq B$ 都属于 $\mathscr{F}$.

以 $\mathscr{N}$ 表示 Ω 的所有这样的子集 N 的类, 对于每个这样的集合 N, 都能找到一个 (自己的) 集合 $B_N \in \mathscr{F}$, 有 $\mathbf{P}(B_N) = 0$, 使得 $N \subseteq B_N$. 令 $\overline{\mathscr{F}}$ (通常也写为 $\mathscr{F}^{\mathbf{P}}$ 或 $\overline{\mathscr{F}}^{\mathbf{P}}$) 为所有形如 $A \cup N$ 的集合的类, 其中 $A \in \mathscr{F}$, 而 $N \in \mathscr{N}$.

证明,

(a) $\overline{\mathscr{F}}$ 是 σ-代数.

(b) 如果 $B \subseteq \Omega$, 并且存在 $\mathscr{F}$ 中的集合 A_1 和 A_2, 使得 $A_1 \subseteq B \subseteq A_2$, 以及 $\mathbf{P}(A_2 \backslash A_1) = 0$, 则有 $B \in \overline{\mathscr{F}}$.

(c) 概率空间 $(\Omega, \overline{\mathscr{F}}, \mathbf{P})$ 是完备的.

§3. 在可测空间上建立概率测度的方法

1. 设 $F(x)=\mathbf{P}(-\infty,x]$. 试证明下列各式的正确性:

$$\mathbf{P}(a,\,b]=F(b)-F(a),\qquad \mathbf{P}(a,\,b)=F(b-)-F(a),$$
$$\mathbf{P}[a,\,b]=F(b)-F(a-),\quad \mathbf{P}[a,\,b)=F(b-)-F(a-),$$
$$\mathbf{P}(\{x\})=F(x)-F(x-),$$

其中 $F(x-)=\lim\limits_{y\uparrow x}F(y)$.

2. 试验证公式 (7) 的正确性.①

3. 证明定理 2.②

4. 证明, $\mathbb{R}$ 上的分布函数 $F=F(x)$ 至多有可数个间断点. 而对于 $\mathbb{R}^n$ 上的分布函数, 可能有怎样的相应结论?

提示: 利用事实 $\{x\in\mathbb{R}:\ F(x)\neq F(x-)\}=\bigcup\limits_{n=1}^{\infty}\{x:\ F(x)-F(x-)\geqslant\frac{1}{n}\}$. 而对于 $\mathbb{R}^n$ $(n\geqslant 2)$ 上的分布函数, 间断点集合具有可数的势的结论已经不再成立. 作为例子, 可观察 δ-测度:

$$\delta_0(A)=\begin{cases}1, & \text{如果 } 0\in A,\\ 0, & \text{如果 } 0\notin A,\end{cases}\qquad A\in\mathscr{B}(\mathbb{R}^n).$$

5. 证明, 如下两个函数都是右连续, 并且对每个变元是上升的, 但却都不是 $\mathbb{R}^2$ 中的 (广义) 分布函数:

$$G(x,\,y)=\begin{cases}1, & x+y\geqslant 0,\\ 0, & x+y<0,\end{cases}$$

$$G(x,\,y)=[x+y],\ \text{即 } x+y \text{ 的整数部分}.$$

①此处所要证明的 (7) 式是:

$$\Delta_{a_1b_1}\cdots\Delta_{a_nb_n}F_n(x_1,\cdots,x_n)=\mathbf{P}(a,\,b],$$

其中 $(a,\,b]=(a_1,\,b_1]\times\cdots\times(a_n,\,b_n]$, 而 $F_n(x_1,\cdots,x_n)$ 是 $\mathbb{R}^n$ 中的分布函数, 且

$$\Delta_{a_ib_i}F_n(x_1,\cdots,x_n)=F_n(x_1,\cdots,x_{i-1},b_i,x_{i+1},\cdots,x_n)-F_n(x_1,\cdots,x_{i-1},a_i,x_{i+1},\cdots,x_n)$$

—— 译者注.

②此处所要证明的定理 2 是: 设 $F=F_n(x_1,\cdots,x_n)$ 是 $\mathbb{R}^n$ 中的分布函数. 则在 $(\mathbb{R}^n,\ \mathscr{B}(\mathbb{R}^n))$ 中存在唯一的概率测度 $\mathbf{P}$, 使得

$$\mathbf{P}(a,\,b]=\Delta_{a_1b_1}\cdots\Delta_{a_nb_n}F_n(x_1,\cdots,x_n)$$

—— 译者注.

6. 设 μ 是某个连续的广义分布函数所导出的勒贝格–斯蒂尔切斯测度. 证明, 如果 A 为至多可数集, 则有 $\mu(A)=0$.

7. 证明, 区间 $[0,1]$ 上的康托尔集 $\mathcal{N}$ 是不可数的完备集 (即为闭的, 同时又稠密于自己的, 或者说没有孤立点的集合), 但是在区间 $[0,1]$ 中却无处稠密, 并且它的勒贝格测度为 0.

8. 设 $(\Omega,\,\mathscr{F},\,\mathbf{P})$ 为概率空间, $\mathscr{A}$ 为 Ω 的子集的代数, 使得 $\mathscr{F}=\sigma(\mathscr{A})$. 试利用适当集合原理, 证明, 对任何 $\varepsilon>0$ 和任何集合 $B\in\mathscr{F}$, 都存在集合 $A_\varepsilon\in\mathscr{A}$, 使得

$$\mathbf{P}(A_\varepsilon\triangle B)\leqslant\varepsilon.$$

提示: 令 $\mathscr{B}=\{B\in\mathscr{F}:\ \forall\,\varepsilon>0,\ \exists\,A\in\mathscr{A}:\ \mathbf{P}(A\triangle B)\leqslant\varepsilon\}$, 证明 $\mathscr{B}$ 是 σ-代数, 从而 $\mathscr{F}\supseteq\mathscr{B}\supseteq\sigma(\mathscr{A})=\mathscr{F}$.

9. 设 $\mathbf{P}$ 为 $(\mathbb{R}^n,\ \mathscr{B}(\mathbb{R}^n))$ 中的概率测度. 证明, 对任何 $\varepsilon>0$ 和 $B\in\mathscr{B}(\mathbb{R}^n)$, 都可以找到紧集 A_1 和开集 A_2, 使得 $A_1\subseteq B\subseteq A_2$, 并且 $\mathbf{P}(A_2\backslash A_1)\leqslant\varepsilon$. (这一事实在定理 3 的证明中曾经用到.)

提示: 考察集合类

$$\mathscr{B}=\left\{\begin{array}{l}B\in\mathscr{B}(\mathbb{R}^n):\ \forall\varepsilon>0,\ \text{存在紧集 } A_1 \text{ 和开集 } A_2,\\ \text{并且 } A_2 \text{ 的闭包 } \overline{A}_2 \text{ 为紧集, 使得 } A_1\subseteq B\subseteq A_2 \text{ 和 } \mathbf{P}(A_2\backslash A_1)\leqslant\varepsilon\end{array}\right\},$$

证明 $\mathscr{B}$ 是 σ-代数.

10. 验证根据公式 $\mathbf{P}_\tau(B)=\mathbf{P}(\mathscr{I}_\tau(B))$ (定理 4, 公式 (21)) 所构造出的测度族 $\{\mathbf{P}_\tau\}$ 的相容性, 其中 $\mathbf{P}$ 为给定的概率测度.

11. 试验证《概率》第一卷第二章第 3 节中的表 2 和表 3 中所列入的 "分布" 都是真正的概率分布.

12. 证明, 第 1 小节注 2 中的类 $\widehat{\mathscr{A}}$ 是 σ-代数.

13. 证明, 第 1 小节注 2 中所引入的集合函数 $\mu(A)$, $A\in\widehat{\mathscr{A}}$, 是测度.

14. 试举例说明, 如果测度 μ_0 在代数 $\mathscr{A}$ 上是有限可加的 (但不是可数可加的), 则它不可能扩展为 $\sigma(\mathscr{A})$ 上的可数可加的测度.

15. 证明, 任何在 Ω 的子集的代数 $\mathscr{A}$ 上给出的有限可加的概率测度都可以扩展为 Ω 的所有子集上的有限可加的概率.

16. 设 $\mathbf{P}$ 是 Ω 的子集的 σ-代数 $\mathscr{F}$ 上的概率测度. 设集合 $C\subseteq\Omega$, 但 $C\notin\mathscr{F}$. 证明, 测度 $\mathbf{P}$ 可以扩展到由 $\mathscr{F}$ 中的集合与 C 所生成的 σ-代数 $\sigma(\mathscr{F},\,C)$ 上 (并保持其可数可加性).

17. 证明, 连续分布函数 F 的支撑 $\operatorname{supp}F$ 是完备集 (定义见第 7 题).

注: $\mathbb{R}$ 上的分布函数 F 的支撑 $\operatorname{supp}F$ 定义为满足如下条件的最小闭集 G: $\mu(\mathbb{R}\backslash G)=0$, 其中 μ 是与分布函数 F 相应的测度; 亦可参阅附录第 2 节.

试举出 $\operatorname{supp} F=\mathbb{R}$, 即以整个实数轴为支撑的离散分布函数 F 的例子.

18. 试证明如下的关于分布函数结构的基本结果 (参阅第 1 小节末尾): 每个分布函数都可以表示为如下形式

$$F=\alpha_1 F_{\mathrm{d}}+\alpha_2 F_{\mathrm{abc}}+\alpha_3 F_{\mathrm{sc}},$$

其中 $\alpha_i \geqslant 0,\ \alpha_1+\alpha_2+\alpha_3=1$,

F_{d} 是离散的分布函数 (在某些点 x_k 处的跃度为 $p_k>0$):

$$F_{\mathrm{d}}=\sum_{\{k:\ x_k \leqslant\ \cdot\}} p_k;$$

F_{abc} 是绝对连续的分布函数:

$$F_{\mathrm{abc}}=\int_{-\infty}^{\cdot} f(t)dt,$$

其中 (密度) $f=f(t)$ 是非负的博雷尔函数, 勒贝格可积, 并且 $\int\limits_{-\infty}^{\infty} f(t)dt=1$;

F_{sc} 是奇异的分布函数, 亦即连续的分布函数, 但是其增长点集合的勒贝格测度为 0.

关于所述的表达式的唯一性问题有何说法?

19. (a) 证明, 每个数 $\omega \in[0,1]$ 都可以表示为 (三进制) 形式

$$\omega=\sum_{n=1}^{\infty} \frac{\omega_n}{3^n},$$

其中 $\omega_n \in\{0,1,2\},\ n \geqslant 1$.

(b) 证明, 如果 ω 有两个表达式 $\sum\limits_{n=1}^{\infty} \frac{\omega_n}{3^n}$ 和 $\sum\limits_{n=1}^{\infty} \frac{\omega_n'}{3^n}$, 并且它们都是无限级数 $\left(\sum\limits_{n=1}^{\infty}|\omega_n|=\infty,\ \sum\limits_{n=1}^{\infty}|\omega_n'|=\infty\right)$, 则有 $\omega_n=\omega_n',\ n \geqslant 1$(无限级数表达式的唯一性).

我们指出, 当 ω 具有有限级数表达式 $\sum\limits_{n=1}^{m} \frac{\omega_n}{3^n}$ 时, 表达方式可以不唯一, 此种场合下可以按照如下方式选择 "典则" 表达方式:

1) 如果 $\omega=\sum\limits_{n=1}^{m-1} \frac{\omega_n}{3^n}+\frac{2}{3^m}$, 则可令

$$\omega_n'=\begin{cases}\omega_n, & n \leqslant m-1, \\ 2, & n=m, \\ 0, & n \geqslant m+1;\end{cases}$$

2) 如果 $\omega=\sum\limits_{n=1}^{m-1} \frac{\omega_n}{3^n}+\frac{1}{3^m}$, 则可令

$$\omega_n'=\begin{cases}\omega_n, & n \leqslant m-1, \\ 0, & n=m, \\ 2, & n \geqslant m+1.\end{cases}$$

(c) 证明, 区间 $[0,1]$ 上的康托尔函数 (康托尔函数是典型的奇异分布函数的例子; 对它的描述见《概率》第一卷第二章第 3 节第 1 小节) 的增长点集合 $\mathcal{N}$ 中的每个 ω 都可以表示为形式:

$$\omega=\sum_{n=1}^{\infty} \frac{\omega_n}{3^n},$$

其中 $\omega_n \in\{0,2\},\ n \geqslant 1$.

注: 有趣的是, 如果考察康托尔集合 $\mathcal{N}$ 中的点 ω 的十进制表达式, 则一共只有 14 个数 (参阅 [118]) 具有有限的级数表达式, 它们是:

$$\frac{1}{4}, \frac{3}{4}, \frac{1}{10}, \frac{3}{10}, \frac{7}{10}, \frac{9}{10}, \frac{1}{40}, \frac{3}{40}, \frac{9}{40}, \frac{13}{40}, \frac{27}{40}, \frac{31}{40}, \frac{37}{40}, \frac{39}{40}.$$

20. 设 $\mathcal{N}$ 是区间 $[0,1]$ 上的康托尔集.

(a) 证明, $\mathcal{N}$ 的势与区间 $[0,1]$ 的势相同.

(b) $\mathcal{N} \oplus \mathcal{N}$ 与 $\mathcal{N} \ominus \mathcal{N}$ 分别等于什么 (其中 $\mathcal{N} \oplus \mathcal{N}=\{\omega+\omega':\ \omega \in \mathcal{N},\ \omega' \in \mathcal{N}\}$, 而 $\mathcal{N} \ominus \mathcal{N}=\{\omega-\omega':\ \omega \in \mathcal{N},\ \omega' \in \mathcal{N}\}$)?

21. 设 C 是 $\mathbb{R}$ 中的闭集. 试构造一个分布函数 F, 使得 $\operatorname{supp} F=C$.

22. 试给出 $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ 中的一个 σ 有限的测度的例子, 使得:

(a) 它不是勒贝格–斯蒂尔切斯测度, 换言之, 对于它不存在非降右连续函数 (广义分布函数) $G=G(x)$, 使得 $\mu((a, b])=G(b)-G(a),\ a<b$.

(b) 它不是局部有限的, 换言之, 对于任何 $x \in \mathbb{R}$, 它在 x 的任何开邻域中的测度都是无限的.

23. 试构造出区间 $[0,1]$ 中的一个集合, 使得它不属于勒贝格集合类 $\overline{\mathscr{B}}([0,1])$.

24. (关于黎曼 ζ 函数欧拉公式的概率证明) 设 $\zeta(\alpha)=\sum\limits_{n=1}^{\infty} \frac{1}{n^\alpha}$ 为黎曼 ζ 函数 $(1<\alpha<\infty)$, 欧拉公式断言, $\zeta(\alpha)$ 满足如下的关系式:

$$[\zeta(\alpha)]^{-1}=\prod_{n=1}^{\infty}\left(1-\frac{1}{p_n^\alpha}\right),$$

其中 $p_1, p_2, \cdots$ 为大于 1 的素数序列. 试证明这一结论.

提示: 令 $\mathbb{N}=\{1,2,\cdots\}$ 为自然数集, 将它的子集的博雷尔 σ-代数记为 $2^{\mathbb{N}}$ (试比较《概率》第一卷第一章第 1 节第 3 小节末尾关于 $N(\mathscr{A})$ 的表示方式). 在 $(\mathbb{N}, 2^{\mathbb{N}})$ 上对子集 $A\subseteq\mathbb{N}$ 定义概率测度 $\mathbf{P}=\mathbf{P}(\cdot)$ 为

$$\mathbf{P}(A)=[\zeta(\alpha)]^{-1}\sum_{n\in A}\frac{1}{n^{\alpha}}.$$

设 $A(p_i)=\{p_i,2p_i,\cdots\}$ 为所有可被素数 p_i 整除的自然数 $n\in\mathbb{N}$ 的集合. 证明,

(a) $\mathbf{P}(A(p_i))=p_i^{-\alpha}$.

(b) 事件 $A(p_1),A(p_2),\cdots$ 相互独立.

(c) $\bigcap\limits_{i=1}^{\infty}\overline{A}(p_i)=\{1\}$.

接下来再考虑如下问题: 由 (b) 可知, 事件 $\overline{A}(p_1),\overline{A}(p_2),\cdots$ 也相互独立. 因此, 由 (c) 和 (a) 可以推出, 一方面

$$\mathbf{P}\left(\bigcap_{i=1}^{\infty}\overline{A}(p_i)\right)=\mathbf{P}(\{1\})=[\zeta(\alpha)]^{-1};$$

另一方面, 又有

$$\mathbf{P}\left(\bigcap_{i=1}^{\infty}\overline{A}(p_i)\right)=\prod_{i=1}^{\infty}[1-\mathbf{P}(A(p_i))]=\prod_{i=1}^{\infty}\left(1-\frac{1}{p_i^{\alpha}}\right),$$

此即为所证.

25. (关于不超过给定数的素数数目的欧拉公式的概率证明) 设 $\varphi(n)$ 为欧拉函数, 即满足条件 $1<p\leqslant n$ 的素数 p 的个数. 试从概率的角度, 证明如下的欧拉公式:

$$\frac{\varphi(n)}{n}=\prod_{p\,|\,n}\left(1-\frac{1}{p}\right)$$

(其中 $p\,|\,n$ 表示 p 可整除 n).

提示: 在集合 $\{1,\cdots,n\}$ 上定义离散的概率分布 $\mathbf{P}(\{k\})=1/n,\ 1\leqslant k\leqslant n$.

对于固定的 n, 令 $A(p)=\{k\leqslant n:\ p\text{整除}k\}$, 并设 $p_1,p_2,\cdots$ 为可以整除 n 的互不相同的素数. 证明

(a) $\mathbf{P}(A(p_i))=p_i^{-1}$.

(b) 事件 $A(p_1),A(p_2),\cdots$ 相互独立.

(c) $\bigcap\limits_{p\,|\,n}\overline{A}(p)$ 就是集合 $\{k\leqslant n:\ k\text{为素数}\}$.

然后再断言

$$\frac{\varphi(n)}{n}=\mathbf{P}\left(\bigcap_{p\,|\,n}\overline{A}(p)\right)=\prod_{p\,|\,n}[1-\mathbf{P}(A(p))]=\prod_{p\,|\,n}\left(1-\frac{1}{p}\right).$$

26. 证明, $(\mathbb{R}^n,\ \mathscr{B}(\mathbb{R}^n))$ 上的勒贝格测度, 在 $n=1$ 时是平移不变的, 而在 $n\geqslant 2$ 时是旋转不变的.

27. 举例说明, 在卡拉泰奥多里定理中关于集合类 $\mathscr{A}$ 是代数的假设, 对于将 $\mathscr{A}$ 上所给定的概率测度 $\mathbf{P}$ 拓广到 σ-代数 $\mathscr{F}=\sigma(\mathscr{A})$ 上的可能性及唯一性来说, 都是本质的. 意即,

(a) 构造一个空间 Ω 和它的子集类 $\mathscr{E}$ 与 $\mathscr{F}$, 其中 $\mathscr{E}$ 不是代数, 而 $\mathscr{F}=\sigma(\mathscr{E})$; 并且在 $\mathscr{E}$ 上给出一个这样的概率测度 $\mathbf{P}$, 它不能拓广到 σ-代数 $\mathscr{F}$ 上.

(b) 构造一个空间 Ω 和它的子集类 $\mathscr{E}$ 与 $\mathscr{F}$, 其中 $\mathscr{F}=\sigma(\mathscr{E})$; 并且在 σ-代数 $\mathscr{F}$ 上给出两个概率测度 $\mathbf{P}$ 与 $\mathbf{Q}$, 它们在 $\mathscr{E}$ 上的局限 (亦即 $\mathbf{P}\,|\,\mathscr{E}$ 与 $\mathbf{Q}\,|\,\mathscr{E}$; 参见第 4 小节) 彼此重合.

提示: (a) 取 $\Omega=\{1,2,3\}$. 令 $\mathscr{E}=\{\varnothing,\ \{1\},\ \{1,2\},\ \{1,3\},\ \Omega\}$, 而 $\mathscr{F}$ 为 Ω 的所有子集所构成的 σ-代数. 再令 $\mathbf{P}(\Omega)=\mathbf{P}(\{1,2\})=\mathbf{P}(\{1,3\})=1,\ \mathbf{P}(\{1\})=1/2,\ \mathbf{P}(\varnothing)=0$.

(b) 只需取 $\Omega=\{1,2,3,4\}$. 并且令 $\mathscr{E}=\{\{1,2\},\ \{1,3\}\}$, 而 $\mathscr{F}$ 为 Ω 的所有子集所构成的 σ-代数. 再令 $\mathbf{P}(\{2\})=\mathbf{P}(\{4\})=1/2,\ \mathbf{Q}(\{2\})=\mathbf{Q}(\{3\})=1/2$.

28. 设 $F=F(x),\ x\in\mathbb{R}$ 为分布函数. 证明, 对于一切 $a\geqslant 0$, 都有

$$\int_{\mathbb{R}}[F(x+a)-F(x)]dx=a.$$

29. 在对于分布函数 $F(x)$ 的密度函数 $f(x)\ (x\in\mathbb{R})$ 的定义中, 仅要求它是非负函数、黎曼可积, 使得 $F(x)=\int\limits_{-\infty}^{x}f(t)dt$, 有时 (例如, 当知道它仅仅黎曼可积, 而非勒贝格可积时) 并不假定它是博雷尔可测的.

试举出一个函数 $f(x)$ 的例子, 它非博雷尔可测, 但却是分布函数 (在所述的意义下) 的密度函数, 并且还能够 (按照公式 $\mu(B)=\int\limits_{-\infty}^{\infty}f(x)I_B(x)dx$) 在数轴 $\mathbb{R}$ 中的博雷尔集合 B 上给出概率测度.

提示: 区间 $[0,1]$ 上所有博雷尔集合的类具有连续统的势 $\mathfrak{c}$, 而所有勒贝格集合的类的势为 $2^{\mathfrak{c}}$ (参阅第 2 节第 32 题). 由此可知, 如果 $\mathcal{N}$ 是区间 $[0,1]$ 上的康托尔集, 则可找到集合 $D\subseteq[1/2,1]\cap\mathcal{N}$, 它不是博雷尔集合, 但却有勒贝格测度 0. 进而再确认, 函数 $f(x)=2I_{[1/2,1]\setminus D}(x)$ 不是博雷尔可测的, 然而黎曼可积, 积分 $\int f(x)I_B(x)dx$ 确定并且给出了区间 $[0,1]$ 中的博雷尔集的概率测度.

30. 试举出实直线 $\mathbb{R}$ 中的两个集合 A 与 B 的例子, 它们都是勒贝格零测集, 然而却有 $A\oplus B=\mathbb{R}$.

31. 设 $\mathscr{A}$ 是 Ω 的子集的某个代数, $\mathscr{F}=\sigma(\mathscr{A})$. 设 μ 是 $\mathscr{F}$ 上的 σ-有限的测度. 证明

(a) 测度 μ 在 $\mathscr{A}$ 上可能不是 σ-有限的.

(b) 如果测度 μ 在 $\mathscr{A}$ 上是 σ-有限的, 则第 8 题中的相应结论亦可成立, 即对一切 $\varepsilon > 0$ 和所有满足条件 $\mu(B) < \infty$ 的 $B \in \mathscr{F}$, 都能找到 $A_\varepsilon \in \mathscr{A}$, 使得 $\mu(A_\varepsilon \triangle B) < \varepsilon$.

(c) 如果测度 μ 在 $\mathscr{A}$ 上不是 σ-有限的, 则性质 (b) 可能不被满足.

32. 证明, $(\mathbb{R}^n,\ \mathscr{B}(\mathbb{R}^n))$ 上的任何概率测度 μ 都是在如下意义下正则的: 对任何 $B \in \mathscr{B}(\mathbb{R}^n)$, 都有

$$\mu(B) = \inf_U \{\mu(U):\ U \supseteq B,\ U \text{ 为开集}\},$$

$$\mu(B) = \sup_F \{\mu(F):\ F \subseteq B,\ F \text{ 为闭集}\}.$$

证明, 对一切博雷尔集 $B \in \mathscr{B}(\mathbb{R}^n)$, 都有

$$\mu(B) = \sup_K \{\mu(K):\ K \subseteq B,\ K \text{ 为紧集}\}.$$

33. (贝特朗悖论) 众所周知的由贝特朗提出的 "经典" 悖论很好地解释了这样的道理: 在解答概率问题时 (尤其是在解答几何概型问题时, 贝特朗悖论就是关于一个几何概型问题的), 应当明晰地描述概率模型, 准确地确定诸如 "侥幸地选取的点", "随机地选取的对象" 之类的语句所表述的意思. (在第一章第 1 节第 17 题的提示中也谈到了这一点.)

贝特朗在他所著的教科书 [8] 中所提出的问题, 以及它的由不同解法所得到的不同结论 (其中包括 "悖论") 都告诉我们: 当概率试验具有无限多个不同的可能结果时, 原则上是不能对某些事件的概率问题给出确定的回答的.

贝特朗所提出的问题如下: 在半径为 r 的圆中 "以随机的方式"(?!) 选取一条两个端点 A 和 B 都位于圆周上的弦 AB. 试问, 这条 "随机选取" 的弦 AB 的长度 $|AB|$ 小于 r 的概率是多少?

我们来陈述关于这个问题的三种不同的可能表述:

(a) 将 "以随机的方式选取弦 AB" 理解为依照圆周上的均匀分布相互独立地选取端点 A 和 B.

试证明, 在关于问题的这种理解之下, 我们所求的概率 $\mathbf{P}_{\rm a}\{|AB| < r\} = 1/3$. (固定点 A, 并考察边长为 r 的圆内接正六边形, 它的一个顶点重合于点 A.)

(b) 每一条弦 AB 都被由圆心向其所作的垂线的垂足 M 所完全决定. 把 "以随机的方式选取弦 AB" 理解为点 M 服从圆内的均匀分布.

试证明, 在关于问题的这种理解之下, 我们所求的概率 $\mathbf{P}_{\rm b}\{|AB| < r\} = 1/4$. (验证事件 $\{|AB| < r\}$ 重合于如下事件: 点 M 位于两个半径分别为 r 和 $r\sqrt{3/4}$ 的同心圆所夹的环中.)

(c) 弦 AB 的长度与其方向无关, 而仅仅取决于它与圆心的距离. 由此自然可以假定它们的方向相同, 例如都平行于水平方向的直径 CD, 而它们与竖直方向的 (与 CD 垂直的) 直径 EF 的交点 M 服从 EF 上的均匀分布.

试证明, 在关于问题的这种理解之下, 所求的概率为

$$\mathbf{P}_{\rm c}\{|AB| < r\} = 1 - \frac{\sqrt{3}}{2} \quad (\approx 0.13).$$

(应当验证事件 $\{|AB| < r\}$ 重合于事件 $\{|OM| > \frac{\sqrt{3}r}{2}\}$, 其中 $|OM|$ 是圆心 O 到点 M 的距离.)

34. (第 33 题续) 上题中针对不同的概率问题所给出的不同解答都是对的. 如果我们以两个参数 ρ 和 θ 来刻画弦, 其中 ρ 是圆心 O 到点 M (由圆心 O 向弦所作的垂线的垂足) 的距离, 而 θ 是弦 AB 与某个固定的方向所形成的夹角. 于是, 如果假定 $r = 1$, 那么对于长度 $|AB| > 0$ 的弦就有 $0 < \rho \leqslant 1,\ 0 \leqslant \theta < 2\pi$.

试证明, 在上题中所考察的三种场合 (a),(b) 与 (c) 之下, (ρ, θ) 的联合分布的密度函数分别为

$$p_{\rm a}(\rho, \theta) = \frac{1}{2\pi^2\sqrt{1-\rho^2}}, \quad p_{\rm b}(\rho, \theta) = \frac{\rho}{\pi}, \quad p_{\rm c}(\rho, \theta) = \frac{1}{2\pi}.$$

(由此即可明白, 并无什么 "悖论" 存在, 因为对 "随机地选择一弦" 一词可以赋予不同的概率意义.)

35. 设 $(X,\ \mathscr{X},\ \mu)$ 为赋予了可数可加的测度 μ 的可测空间 $(X,\ \mathscr{X})$(参阅第 1 节定义 5 和定义 6).

将可测集 A 称为 (相对于测度 μ 的) *原子*或者 *μ-原子*, 如果 $\mu(A) > 0$, 并且对于任何可测集 B, 都或有 $\mu(A \cap B) = 0$, 或有 $\mu(A \backslash B) = 0$. 测度 μ 称为*原子的*, 如果任何 μ 测度为正的可测集都包含着原子.

测度 μ 称为*非原子的*, 如果不存在 μ-原子.

测度 μ 称为*弥漫的*, 如果单点集都是可测的, 并且 μ 测度都为 0.

试分别给出原子的、非原子的、弥漫的测度的例子, 并给出同时为原子的和弥漫的测度的例子.

证明, 原子的与非原子的测度之和可能是原子的.

36. 设 $\mathbf{P}$ 与 $\widetilde{\mathbf{P}}$ 是 $(\Omega, \mathscr{F})$ 中的两个概率测度, 现知对任何满足不等式 $\mathbf{P}(A) \leqslant 1/2$ 的 $\mathscr{F}$ 中的集合 A, 都有 $\mathbf{P}(A) = \widetilde{\mathbf{P}}(A)$. 证明, 对一切 $A \in \mathscr{F}$, 都有 $\mathbf{P}(A) = \widetilde{\mathbf{P}}(A)$.

§4. 随机变量 I

1. 证明, 随机变量 ξ 连续, 当且仅当, 对任何 $x \in \mathbb{R}$, 都有 $\mathbf{P}\{\xi = x\} = 0$.

2. 如果 $|\xi|$ 是 $\mathscr{F}$ 可测的, 那么 ξ 是否也一定是 $\mathscr{F}$ 可测的?

3. 证明, 函数 $x^n,\ x^+ = \max(x, 0),\ x^- = -\min(x, 0),\ |x| = x^+ + x^-$ 都是博雷尔的. 证明进一步的结论: 任何连续函数 $f = f(x),\ x \in \mathbb{R}$ 都是博雷尔的.

提示: 对任何 $\alpha \in \mathbb{R}$, 考察开集 $\{\omega \in \mathbb{R},\ f(\omega) < \alpha\}$, 并利用第 2 节第 7 题的结论.

4. 证明, 如果 ξ 与 η 都是 $\mathscr{F}$ 可测的, 则 $\{\omega:\ \xi(\omega) = \eta(\omega)\} \in \mathscr{F}$.

5. 设 ξ 与 η 是 $(\Omega, \mathscr{F})$ 上的两个随机变量, 集合 $A \in \mathscr{F}$. 证明, 如下的函数也是随机变量:

$$\zeta(\omega) = \xi(\omega) I_A + \eta(\omega) I_{\overline{A}}.$$

6. 设 $\xi_1, \cdots, \xi_n$ 为随机变量, 而 $\varphi(x_1, \cdots, x_n)$ 为博雷尔函数. 证明, 函数 $\varphi(\xi_1(\omega), \cdots, \xi_n(\omega))$ 也是随机变量.

提示: 首先证明, 如下的映射是 $\mathscr{F}/\mathscr{B}(\mathbb{R}^n)$ 可测的:

$$\omega \mapsto (\xi_1(\omega), \cdots, \xi_n(\omega)) \in \mathbb{R}^n.$$

再证明 $\omega \mapsto \varphi(\xi_1(\omega), \cdots, \xi_n(\omega))$ 是复合可测映射.

7. 设 ξ 与 η 是两个取值于 $1, \cdots, N$ 的随机变量. 假设 $\mathscr{F}_\xi = \mathscr{F}_\eta$. 证明, 存在数 $(1, \cdots, N)$ 的一个排列 $(i_1, \cdots, i_N)$, 使得对于每个 $j = 1, \cdots, N$, 都有集合 $\{\omega:\ \xi = j\}$ 与集合 $\{\omega:\ \eta = i_j\}$ 相互重合.

提示: 利用定理 3, 根据该定理, 可以找到函数 φ 与 ψ, 使得 $\xi = \varphi(\eta)$ 与 $\eta = \psi(\xi)$. 再证明, $i_j = \psi(j)$ 就是所求的排列.

8. 举例说明, 存在这样的随机变量 ξ, 它的分布函数具有密度函数 $f(x)$, 但是 $\lim\limits_{x\to\infty} f(x)$ 不存在, 由此 $f(x)$ 在无穷远处并不零化.

9. 设 ξ 与 η 是两个有界的随机变量 ($|\xi| \leqslant c_1,\ |\eta| \leqslant c_2$). 证明, 如果对于一切 $m, n \geqslant 1$, 都有

$$\mathbf{E}\xi^m\eta^n = \mathbf{E}\xi^m \cdot \mathbf{E}\eta^n,$$

则 ξ 与 η 相互独立.

10. 设 ξ 与 η 是两个随机变量, 它们的分布函数 F_ξ 与 F_η 相同. 证明, 如果对于 $x \in \mathbb{R}$, 有 $\{\omega:\ \xi(\omega) = x\} \neq \varnothing$, 则存在 $y \in \mathbb{R}$, 使得 $\{\omega:\ \xi(\omega) = x\} = \{\omega:\ \eta(\omega) = y\}$.

11. 设 E 是 $\mathbb{R}$ 的至多可数的子集, ξ 为 $\Omega \to E$ 的映射. 证明, ξ 是 $(\Omega, \mathscr{F})$ 上的随机变量, 当且仅当, 对每个 $x \in E$, 都有 $\{\omega:\ \xi(\omega) = x\} \in \mathscr{F}$.

12. 设 ξ 是随机变量, 有 $\mathbf{P}\{\xi \neq 0\} > 0$. 假设对于某两个数 a 和 b, 随机变量 $a\xi$ 与 $b\xi$ 同分布 ($F_{a\xi}(x) = F_{b\xi}(x),\ x \in \mathbb{R}$). 试问, 是否必有 $a = b$? 如果假设 $a \geqslant 0,\ b \geqslant 0$, 那么命题自身是否成立?

13. 设 $(\Omega, \mathscr{F}, \mathbf{P})$ 是概率空间, 而 $(\Omega, \overline{\mathscr{F}}^{\mathbf{P}}, \mathbf{P})$ 是 $(\Omega, \mathscr{F}, \mathbf{P})$ 对测度 $\mathbf{P}$ 的完备化 (参阅本书第 2 节第 34 题, 亦可见《概率》第一卷第二章第 3 节第 1 小节的注 1).

设 $\overline{\xi} = \overline{\xi}(\omega)$ 是 $(\Omega, \overline{\mathscr{F}}^{\mathbf{P}}, \mathbf{P})$ 上的随机变量. 证明, 存在 $(\Omega, \mathscr{F}, \mathbf{P})$ 上的随机变量 $\xi = \xi(\omega)$, 使得 $\mathbf{P}\{\xi \neq \overline{\xi}\} = 0$. 换言之, $\overline{\xi}$ 与 ξ 仅在某个零概集上不相等.

14. 设 ξ 是随机变量, B 是 $\mathbb{R}$ 中的博雷尔集. 证明

$$\sigma(\xi I(\xi \in B)) = \xi^{-1}(B) \cap \sigma(\xi).$$

15. 设 $\xi_1, \xi_2, \cdots$ 为相互独立的区间 $[0,1]$ 上均匀分布的随机变量. 令 $A(\omega)$ 为 $[0,1]$ 上的由值 $\xi_1(\omega), \xi_2(\omega), \cdots$ 所形成的集合, 其中 $\omega \in \Omega$. 证明, 对于几乎一切 $\omega \in \Omega$, 集合 $A(\omega)$ 都在区间 $[0,1]$ 中稠密.

16. 设 $\xi_1, \xi_2, \cdots$ 为相互独立的伯努利随机变量, 有 $\mathbf{P}\{\xi_k = 1\} = \mathbf{P}\{\xi_k = -1\} = 1/2,\ k \geqslant 1$. 我们来考察随机游动 $\{S_n\}_{n\geqslant 0}$, 其中 $S_0 = 0,\ S_n = \xi_1 + \cdots + \xi_n,\ n \geqslant 1$.

令 $\sigma_0 = \inf\{n > 0:\ S_n = 0\}$ 为 0 时刻之后首次回归 0 点的时刻 (如同惯例, 当相应的集合 $\{\cdot\}$ 为空集时, 令 $\sigma_0 = \infty$).

证明

$$\mathbf{P}\{\sigma_0 > 2n\} = C_{2n}^n \left(\frac{1}{2}\right)^{2n} \quad 而 \quad \mathbf{P}\{\sigma_0 = 2n\} = \frac{1}{2n-1} C_{2n}^n \left(\frac{1}{2}\right)^{2n}.$$

再运用斯特林公式, 验证, 当 n 充分大时, 有

$$\mathbf{P}\{\sigma_0 > 2n\} \sim \frac{1}{\sqrt{\pi n}} \quad 而 \quad \mathbf{P}\{\sigma_0 = 2n\} \sim \frac{1}{2\sqrt{\pi} n^{3/2}}$$

(试比较第一章第 10 节中关于 u_{2k} 和 f_{2k} 的表达式). (由所得到的公式推知 $\mathbf{P}\{\sigma_0 < \infty\} = 1$, 并且 $\mathbf{E}\,\sigma_0^\alpha < \infty$ 当且仅当 $\alpha < 1/2$; 试比较第一章第 9 节中的结果.)

17. 在上一题中的条件下, 令 $\sigma_k = \inf\{n \geqslant 1:\ S_n = k\},\ k = 1, 2, \cdots$. 证明

$$\mathbf{P}\{\sigma_k = n\} = \frac{k}{n} \mathbf{P}\{S_n = k\},$$

这就意味着有

$$\mathbf{P}\{\sigma_k = n\} = \frac{k}{n} C_n^{\frac{n+k}{2}} \left(\frac{1}{2}\right)^n.$$

18. 设 $\xi = \xi(\omega)$ 为非退化随机变量, 使得对于常数 $a > 0$ 和 b, 随机变量 $a\xi + b$ 的分布与 ξ 的分布相同. 证明, 此时必有 $a = 1$ 和 $b = 0$.

19. 设 ξ_1 与 ξ_2 为两个可交换的随机变量 (意即 (ξ_1, ξ_2) 与 (ξ_2, ξ_1) 同分布). 证明, 如果 $f = f(x)$ 和 $g = g(x)$ 都是非负的非降函数, 则有

$$\mathbf{E} f(\xi_1) g(\xi_1) \geqslant \mathbf{E} f(\xi_1) g(\xi_2).$$

20. (针对定理 2) 设 $\xi_1, \xi_2, \cdots$ 为 (取值于 $\mathbb{R}$) 的随机变量序列. 证明

$$B = \{\omega: \lim \xi_n(\omega) \text{ 存在且有限}\} \in \mathscr{F}.$$

提示: 利用 B 的如下表达方式:

$$B = \{\underline{\lim}\, \xi_n > -\infty\} \cap \{\overline{\lim}\, \xi_n < \infty\} \cap \{\overline{\lim}\, \xi_n - \underline{\lim}\, \xi_n = 0\}.$$

21. 设 $\xi_1, \xi_2, \cdots$ 为独立随机变量序列, 服从共同的非退化的连续分布. 设 $A_1, A_2, \cdots$ 为一列事件, 其中 $A_1 = \Omega$, 而对于 $n \geqslant 2$, 有

$$A_n = \{\xi_n > \xi_m \text{ 对一切 } m < n\}.$$

(事件 A_n 可以视为在时刻 n 产生 "新记录" 的事件.) 证明, 事件 $A_1, A_2, \cdots$ 相互独立, 并且 $\mathbf{P}(A_n) = 1/n,\ n \geqslant 1$.

22. 设 ξ 与 η 是随机变量, 其中 η 的分布律 $\mathrm{Law}(\eta)$ 绝对连续 (即其分布函数 F_η 是绝对连续函数). 证明,

(a) 如果 ξ 与 η 相互独立, 则分布律 $\mathrm{Law}(\xi + \eta)$ 也是绝对连续的.

(b) 如果 ξ 与 η 不相互独立, 则分布律 $\mathrm{Law}(\xi + \eta)$ 可能不是绝对连续的.

23. 设 ξ 与 η 是相互独立的随机变量, 其中 ξ 为离散的, 而 η 为奇异的 (即 F_ξ 是离散的分布函数, F_η 是奇异连续的分布函数). 证明, 随机变量 $\xi + \eta$ 的概率分布 $F_{\xi+\eta}$ 是奇异的.

24. 设 $(\Omega, \mathscr{F})$ 是可测空间, 并且 σ-代数 $\mathscr{F}$ 是某个 (可数) 分割 $\mathscr{D} = \{D_1, D_2, \cdots\}$(参阅《概率》第一卷第二章第 2 节第 1 小节) 所可数生成的. 证明, σ-代数 $\mathscr{F}$ 重合于如下的随机变量所生成的 σ-代数 $\mathscr{F}_X$:

$$X(\omega) = \sum_{n=1}^{\infty} \frac{\varphi(I_{D_n}(\omega))}{10^n},$$

其中 $\varphi(0) = 3,\ \varphi(1) = 5$.

25. (a) 设随机变量 X 具有对称的分布 (即 $\mathrm{Law}\,(X) = \mathrm{Law}\,(-X)$). 证明,

$$\mathrm{Law}(X) = \mathrm{Law}(\xi Y),$$

其中 ξ 与 Y 为相互独立的随机变量, 有 $\mathbf{P}\{\xi = 1\} = \mathbf{P}\{\xi = -1\} = 1/2$, 而 $\mathrm{Law}\,(Y) = \mathrm{Law}(|X|)$.

(b) 设 ξ 与 Y 为相互独立的随机变量, 有 $\mathbf{P}\{\xi = 1\} = \mathbf{P}\{\xi = -1\} = 1/2$. 证明, ξ 与 ξY 相互独立, 当且仅当, Y 具有对称的分布 (即 $\mathrm{Law}\,(Y) = \mathrm{Law}\,(-Y)$).

26. 设随机变量 X 取两个值 x_1 和 x_2(并且 $x_1 \neq x_2$), 随机变量 Y 取两个值 y_1 和 y_2(并且 $y_1 \neq y_2$). 证明, 如果 $\mathbf{cov}(X, Y) = 0$, 则 X 与 Y 相互独立.

27. 设 $\xi_1, \xi_2, \cdots$ 为独立随机变量序列, 均服从区间 $[0, 1]$ 上的均匀分布, 并且对于 $0 < x < 1$, 有

$$\tau(x) = \min\{n \geqslant 1:\ \xi_1 + \cdots + \xi_n > x\}.$$

证明, $\mathbf{P}\{\tau(x) > n\} = x^n/n!,\ n \geqslant 1$.

28. 设 X_1, X_2, X_3 为独立同分布随机变量, 具有指数密度 $f(x) = e^{-x} I(x > 0)$. 定义

$$Y_1 = \frac{X_1}{X_1 + X_2}, \quad Y_2 = \frac{X_1 + X_2}{X_1 + X_2 + X_3}, \quad Y_3 = X_1 + X_2 + X_3.$$

证明, 随机变量 Y_1, Y_2, Y_3 相互独立.

29. 设 X_1 与 X_2 为相互独立的随机变量, 分别服从自由度为 r_1 和 r_2 的 χ^2 分布 (参阅第 8 节 (34) 式和第 3 节表 3). 证明, 随机变量 $Y_1 = X_1/X_2$ 与 $Y_2 = X_1 + X_2$ 相互独立 (试比较第 13 节中的 39 和 42 题结论).

§5. 随机元

1. 设 $\xi_1, \cdots, \xi_n$ 为离散随机变量. 证明, 它们相互独立, 当且仅当对任何 $x_1, \cdots, x_n$, 都有

$$\mathbf{P}\{\xi_1 = x_1, \cdots, \xi_n = x_n\} = \prod_{i=1}^{n} \mathbf{P}\{\xi_i = x_i\}.$$

2. 试证明, 一切随机函数 $X(\omega) = (\xi_t(\omega))_{t \in T}$ 都是 (定义 3 意义下的) 随机过程, 反之亦然.

提示: 如果 $X = X(\omega)$ 是 $\mathscr{F}/\mathscr{B}(\mathbb{R}^T)$ 可测函数, 则对 $t \in T$ 和 $B \in \mathscr{B}(\mathbb{R})$, 都有

$$\{\omega:\ \xi_t(\omega) \in B\} = \{\omega:\ X(\omega) \in C\} \in \mathscr{F},$$

其中 $C = \{x \in \mathbb{R}^T:\ x_t \in B\}$. 反过来, 只需考察形如 $\{x:\ x_{t_1} \in B_1, \cdots, x_{t_n} \in B_n\}$ 的集合 $C \in \mathscr{B}(\mathbb{R}^T)$, 其中 $B_1, \cdots, B_n \in \mathscr{B}(\mathbb{R})$. 该种集合显然属于 $\mathscr{F}$.

3. 设 $X_1, \cdots, X_n$ 分别为取值于 $(E_1, \mathscr{E}_1), \cdots, (E_n, \mathscr{E}_n)$ 的随机元. 又设 $(E_1', \mathscr{E}_1'), \cdots, (E_n', \mathscr{E}_n')$ 为 n 个可测空间, 而 $g_1, \cdots, g_n$ 分别为 $\mathscr{E}_1/\mathscr{E}_1', \cdots, \mathscr{E}_n/\mathscr{E}_n'$ 可测的函数. 证明, 如果 $X_1, \cdots, X_n$ 相互独立, 则随机元 $g_1 \circ X_1, \cdots, g_n \circ X_n$ 亦相互独立, 其中 $g_i \circ X_i = g_i(X_i),\ i = 1, \cdots, n$.

提示: 只需证明, 对任何 $B_i \in \mathscr{E}_i'$ ①, $i = 1, \cdots, n$, 都有

$$\mathbf{P}\{g_1(X_1) \in B_1, \cdots, g_n(X_n) \in B_n\} = \mathbf{P}\{X_1 \in g_1^{-1}(B_1), \cdots, X_n \in g_n^{-1}(B_n)\}.$$

① 原书为: 对任何 $B_i \in \mathscr{E}_i,\ i = 1, \cdots, n$ —— 译者注.

4. 设 $X_1, X_2, \cdots$ 为可交换随机变量的无限序列 (亦即, k 个脚标不同的随机变量, 例如 $X_{i_1}, \cdots, X_{i_k}$ 的联合分布仅仅与 k 有关, 而与 $i_1, \cdots, i_k$ 的具体选取无关, 试比较第 1 节第 11 题.) 证明, 如果 $\mathbf{E}X_n^2 < \infty$, $n \geqslant 1$, 则 $\mathbf{cov}(X_1, X_2) \geqslant 0$.

提示: 利用可交换性假设, 可以用前两个矩和协方差表示 $\mathbf{D}\left(\sum_{i=1}^n X_i\right)$, 再在其中令 $n \to \infty$, 考察该极限.

5. 设 $\xi_1, \cdots, \xi_m$ 与 $\eta_1, \cdots, \eta_n$ 为随机变量, 记随机向量 $X = (\xi_1, \cdots, \xi_m)$ 与 $Y = (\eta_1, \cdots, \eta_n)$. 假设如下各条件成立:

(i) 随机变量 $\xi_1, \cdots, \xi_m$ 相互独立.

(ii) 随机变量 $\eta_1, \cdots, \eta_n$ 相互独立.

(iii) 随机向量 X 与 Y, 作为分别取值于 $\mathbb{R}^m$ 与 $\mathbb{R}^n$ 的随机元, 也相互独立.

证明, 此时随机变量 $\xi_1, \cdots, \xi_m, \eta_1, \cdots, \eta_n$ 相互独立.

6. 给定随机向量 $X = (\xi_1, \cdots, \xi_m)$ 与 $Y = (\eta_1, \cdots, \eta_n)$, 现知随机变量 $\xi_1, \cdots, \xi_m, \eta_1, \cdots, \eta_n$ 是相互独立的.

(a) 证明, 随机向量 X 与 Y 作为随机元是相互独立的 (试比较上述第 5 题).

(b) 令 $f: \mathbb{R}^m \to \mathbb{R}$, $g: \mathbb{R}^n \to \mathbb{R}$ 为博雷尔函数. 证明, 随机变量 $f(\xi_1, \cdots, \xi_m)$ 与 $g(\eta_1, \cdots, \eta_n)$ 相互独立.

7. 设 $(\Omega, \mathscr{F})$ 为可测空间, 而 $(E, \mathscr{E}, \rho)$ 为以 ρ 为距离的距离空间, $\mathscr{E}$ 是由其中的所有开集所生成的 σ-代数 (参阅第 2 节). 设 $X_1(\omega), X_2(\omega), \cdots$ 为 $\mathscr{F}/\mathscr{E}$ 可测的函数序列 (随机元), 使得对一切 $\omega \in \Omega$, 都存在极限

$$X(\omega) = \lim_{n\to\infty} X_n(\omega).$$

证明, 极限函数 $X(\omega)$ 亦是 $\mathscr{F}/\mathscr{E}$ 可测的.

8. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布随机变量序列, $\mathscr{F}_n = \sigma(\xi_1, \cdots, \xi_n)$, $n \geqslant 1$, 而 $\tau(\omega)$ 为 (相对于 $(\mathscr{F}_n)_{n\geqslant 1}$ 的) 停时. 令

$$\eta_n(\omega) = \xi_{n+\tau(\omega)}(\omega).$$

证明, 序列 $(\eta_1, \eta_2, \cdots)$ 与序列 $(\xi_1, \xi_2, \cdots)$ 同分布.

§6. 勒贝格积分. 数学期望

1. 证明 (6) 式.

提示: 设 S 为非负简单函数 s 的集合. 如果 $s \in \{s \in S: s \leqslant \xi\}$, 而 $(\xi_n)_{n\geqslant 1}$ 是简单随机变量序列, 有 $\xi_n \uparrow \xi$, 则有 $\max(\xi_n, s) \uparrow \xi$ 与 $\mathbf{E}s \leqslant \mathbf{E}\max(\xi_n, s)$, 由此可知 $\mathbf{E}s \leqslant \mathbf{E}\xi$ 和 $\sup\limits_{\{s\in S:\ s\leqslant\xi\}} \mathbf{E}s \leqslant \mathbf{E}\xi$. 反过来的不等式可以由 $\mathbf{E}\xi$ 的结构直接得到.

2. 试证明关于性质 $\mathbf{E}$ 的如下推广: 设 ξ 与 η 是随机变量, 对于它们, $\mathbf{E}\xi$ 与 $\mathbf{E}\eta$ 是确定的, 而且表达式 $\mathbf{E}\xi + \mathbf{E}\eta$ 也有意义 (即不具有形式 $\infty - \infty$ 或 $-\infty + \infty$), 则有

$$\mathbf{E}(\xi + \eta) = \mathbf{E}\xi + \mathbf{E}\eta.$$

提示: 正如证明性质 $\mathbf{E}$ 本身那样, 需要考察由表达式 $\xi = \xi^+ - \xi^-$ 和 $\eta = \eta^+ - \eta^-$ 所引发的各种可能情况. 例如, 如果 $\mathbf{E}\xi^+ = \infty$, 则 (在相应的条件下) 证明 (或反证) 亦有 $\mathbf{E}(\xi + \eta)^+ = \infty$.

3. 试推广性质 $\mathbf{G}$, 即证明, 如果 $\xi = \eta$ (a.s.), 而 $\mathbf{E}\xi$ 存在, 则 $\mathbf{E}\eta$ 也存在, 并且有 $\mathbf{E}\eta = \mathbf{E}\xi$.

4. 设 ξ 为广义随机变量, μ 为 σ-有限测度, 有 $\int_\Omega |\xi| d\mu < \infty$. 证明, 此时必有 $|\xi| < \infty$ (μ-a.s.). (试比较性质 $\mathbf{J}$.)

5. 设 μ 为 σ-有限测度, ξ 与 η 均为广义随机变量, 使得 $\int \xi d\mu$ 与 $\int \eta d\mu$ 都有定义. 那么, 只要对所有 $A \in \mathscr{F}$ 都有不等式 $\int_A \xi d\mu \leqslant \int_A \eta d\mu$ 成立, 就有 $\xi \leqslant \eta$ (μ-a.s.). (试比较性质 $\mathbf{I}$.)

6. 设 ξ 与 η 为相互独立的非负随机变量. 证明, $\mathbf{E}\xi\eta = \mathbf{E}\xi \cdot \mathbf{E}\eta$.

提示: 现将 ξ 与 η 换为这样的简单随机变量 ξ_n 与 η_n, 其中 $\xi_n \uparrow \xi$, $\eta_n \uparrow \eta$. 由定理 6 知 $\mathbf{E}\xi_n\eta_n = \mathbf{E}\xi_n \cdot \mathbf{E}\eta_n$. 再利用单调收敛定理即可.

7. 利用法图引理, 证明

$$\mathbf{P}\left(\underline{\lim} A_n\right) \leqslant \underline{\lim}\, \mathbf{P}\left(A_n\right), \quad \mathbf{P}\left(\overline{\lim} A_n\right) \geqslant \overline{\lim}\, \mathbf{P}\left(A_n\right).$$

8. 举例说明, 在控制收敛定理中, 条件 "$|\xi_n| \leqslant \eta$, $\mathbf{E}\eta < \infty$", 一般来说, 是不能削弱的.

提示: 令 $\Omega = [0, 1]$, $\mathscr{F} = \mathscr{B}([0, 1])$, $\mathbf{P}$ 为勒贝格测度. 考察随机变量 $\xi_n(\omega) = -nI(\omega \leqslant 1/n)$, $n \geqslant 1$.

9. 举例说明, 在法图引理中, 条件 "$|\xi_n| \leqslant \eta$, $\mathbf{E}\eta > -\infty$", 一般来说, 不能取消.

10. 证明法图引理的如下表述方式的正确性: 如果随机变量族 $\{\xi_n^+, n \geqslant 1\}$ 一致可积, 则有

$$\overline{\lim}\, \mathbf{E}\, \xi_n \leqslant \mathbf{E}\, \overline{\lim}\, \xi_n.$$

提示: 利用性质: 对任给的 $\varepsilon > 0$, 可以找到 $c > 0$, 使得对所有 $n \geqslant 1$, 都有 $\mathbf{E}\xi_n I(\xi_n > c) < \varepsilon$.

11. 定义在区间 $[0,1]$ 上的狄利克雷函数

$$d(x)=\begin{cases}1, & x \text{ 为有理数},\\ 0, & x \text{ 为无理数}\end{cases}$$

是勒贝格可积的, 但不是黎曼可积的. 为什么?

12. 举例说明, 存在定义在区间 $[0,1]$ 上的黎曼可积的函数序列 $\{(f_n)_{n\geqslant 1}\}$, 其中 $|f_n|\leqslant 1$, 并且依勒贝格测度几乎处处有 $f_n\to f$, 但是 f 却非黎曼可积.

提示: 考察函数列 $f_n(x)=\sum\limits_{i=1}^{n} I_{\{q_i\}}(x)$, 其中 $\{q_1,q_2,\cdots\}$ 是区间 $[0,1]$ 中的有理数集.

13. 设 $\{a_{ij};\ i,j\geqslant 1\}$ 为实数族, 有 $\sum\limits_{i,j}|a_{ij}|<\infty$. 试由富比尼定理推出

$$\sum_{i,j}a_{ij}=\sum_i\left(\sum_j a_{ij}\right)=\sum_j\left(\sum_i a_{ij}\right). \qquad (*)$$

提示: 任取一列正数 $p_1,p_2,\cdots$, 使得 $\sum\limits_{i=1}^{\infty}p_i=1$, 并且在 $\Omega=\mathbb{N}=\{1,2,\cdots\}$ 上定义概率测度 $\mathbf{P}$ 为 $\mathbf{P}(A)=\sum\limits_{i\in A}p_i$. 然后定义函数 $f(i,j)=\frac{a_{ij}}{p_ip_j}$, 并指出

$$\int_{\Omega\times\Omega}|f(\omega_1,\omega_2)|d(\mathbf{P}\times\mathbf{P})=\sum_{i,j}|f(i,j)|p_ip_j=\sum_{i,j}|a_{ij}|<\infty.$$

最后再利用富比尼定理.

14. 举例说明, 存在这样的实数族 $\{a_{ij};\ i,j\geqslant 1\}$, 有 $\sum\limits_{i,j}|a_{ij}|=\infty$, 使得第 13 题中的 $(*)$ 式不成立.

提示: 考察

$$a_{ij}=\begin{cases}0, & i=j,\\ (i-j)^{-3}, & i\neq j.\end{cases}$$

15. 由简单函数出发, 利用关于勒贝格积分的积分号下取极限的定理, 证明如下的关于积分换元法结论的正确性.

设 $h=h(y)$ 为区间 $[a,b]$ 上的非降的连续可微函数, 而 $f(x)$ 是区间 $[h(a),h(b)]$ 上的 (按勒贝格测度) 可积的函数. 则函数 $f(h(y))h'(y)$ 在区间 $[a,b]$ 上可积, 并且

$$\int_{h(a)}^{h(b)}f(x)dx=\int_a^b f(h(y))h'(y)dy.$$

提示: 首先证明该式对博雷尔集合的示性函数及其线性组合成立. 再利用单调收敛定理, 证明该式对非负 (可积) 函数 $f(x)$ 成立, 最后利用表达式 $f=f^+-f^-$, 证明其对任意 (可积) 函数 $f(x)$ 成立.

16. 证明 (70) 式.

提示: 令 $\widetilde{\xi}=-\xi$, 那么它的分布函数为 $\widetilde{F}(x)=1-F((-x)-)$. 从而 $\int\limits_{-\infty}^{0}|x|^n dF(x)=\int\limits_{0}^{\infty}|x|^n d\widetilde{F}(x)$. 接着再利用 (69) 式.

17. 设 $\xi,\xi_1,\xi_2,\cdots$ 为非负随机变量序列, 对任何 $\varepsilon>0$, 都有 $P(|\xi_n-\xi|>\varepsilon)\to 0,\ n\to\infty$ (即有依概率收敛 $\xi_n\xrightarrow{\mathbf{P}}\xi$; 参阅第 10 节).

(a) 作为定理 5 的推广, 试证明, 在假设 $\mathbf{E}\xi_n<\infty,\ n\geqslant 1$ 之下, 有 $\mathbf{E}\xi_n\to\mathbf{E}\xi<\infty$ 成立的充分必要条件是 $\{\xi_n,\ n\geqslant 1\}$ 一致可积 (换言之, 如果将以概率 1 收敛换为依概率收敛, 则定理 5 的结论仍然成立).

(b) 设所有随机变量 $\xi,\xi_1,\xi_2,\cdots$ 都是可积的 (即 $\mathbf{E}\xi<\infty,\ \mathbf{E}\xi_n<\infty,\ n\geqslant 1$). 证明

$$\mathbf{E}\xi_n\to\mathbf{E}\xi \quad\Rightarrow\quad \mathbf{E}|\xi_n-\xi|\to 0.$$

提示: (a) 充分性可由定理 4 和第 10 节第 1 题推出. 必要性可以类似于定理 5 的证明, 只需将那里的以概率 1 收敛 换为依概率收敛 (当然还应参考第 10 节第 1 题).

(b) 对任何 $c>0$, 都有

$$\mathbf{E}|\xi-\xi_n|\leqslant\mathbf{E}|\xi-(\xi\wedge c)|+\mathbf{E}|(\xi\wedge c)-(\xi_n\wedge c)|+\mathbf{E}|(\xi_n\wedge c)-\xi_n|.$$

给定 $\varepsilon>0$, 选 $c>0$, 使 $\mathbf{E}|\xi-(\xi\wedge c)|<\varepsilon$. (在题中的条件下) 可以断言, 对于充分大的 n, 有 $\mathbf{E}|(\xi\wedge c)-(\xi_n\wedge c)|\leqslant\varepsilon$ 和 $\mathbf{E}|(\xi_n\wedge c)-\xi_n|\leqslant 3\varepsilon$. 从而对于一切充分大的 n, 都有 $\mathbf{E}|\xi-\xi_n|\leqslant 5\varepsilon$.

18. 设 ξ 为可积随机变量 ($\mathbf{E}|\xi|<\infty$).

(a) 证明, 对一切 $\varepsilon>0$, 都存在 $\delta>0$, 使得对任何 $A\in\mathscr{F}$, 只要 $\mathbf{P}(A)<\delta$, 就都有 $\mathbf{E}I_A|\xi|<\varepsilon$ (“勒贝格积分的绝对连续性”).

(b) 试由 (a) 推出: 如果 $(A_n)_{n\geqslant 1}$ 是一列事件, 有 $\lim\limits_{n\to\infty}\mathbf{P}(A_n)=0$, 则有 $\mathbf{E}(\xi I(A_n))\to 0,\ n\to\infty$.

提示: 利用《概率》第一卷第二章第 6 节引理 2.

注: 试比较此处的断言 (b) 和定理 3 中的断言.

19. 设 ξ,η,ζ 与 $\xi_n,\eta_n,\zeta_n,\ n\geqslant 1$, 为满足如下条件的随机变量 (参阅第 17 题中关于依概率收敛 “$\xrightarrow{\mathbf{P}}$” 的定义):

$$\xi_n\xrightarrow{\mathbf{P}}\xi,\quad \eta_n\xrightarrow{\mathbf{P}}\eta,\quad \zeta_n\xrightarrow{\mathbf{P}}\zeta,\quad \eta_n\leqslant\xi_n\leqslant\zeta_n,\quad n\geqslant 1,$$
$$\mathbf{E}\zeta_n\to\mathbf{E}\zeta\quad \mathbf{E}\eta_n\to\mathbf{E}\eta,$$

并且 $\mathbf{E}\xi,\ \mathbf{E}\eta,\ \mathbf{E}\zeta$ 均有限. 证明, 此时有如下各断言 (普拉特 (Pratt) 引理) 成立:

(a) $\mathbf{E}\xi_n\to\mathbf{E}\xi$.

(b) 如果还有条件 $\eta_n \leqslant 0 \leqslant \zeta_n$ 成立, 那么就还有 $\mathbf{E}|\xi_n - \xi| \to 0$.

并由此推出, 如果 $\xi_n \xrightarrow{\mathbf{P}} \xi$, $\mathbf{E}|\xi_n| \to \mathbf{E}|\xi|$ 及 $\mathbf{E}|\xi| < \infty$, 则 $\mathbf{E}|\xi_n - \xi| \to 0$. (试比较第 17 题的断言 (b).)

举例说明, 如果去掉 (b) 中的补充条件, 就可能会有 $\mathbf{E}|\xi_n - \xi| \not\to 0$.

提示: 如果引入辅助随机变量

$$\widetilde{\eta}_n = 0, \quad \widetilde{\xi}_n = \xi_n - \eta_n, \quad \widetilde{\zeta}_n = \zeta_n - \eta_n \quad 与 \quad \widetilde{\eta} = 0, \quad \widetilde{\xi} = \xi - \eta, \quad \widetilde{\zeta} = \zeta - \eta,$$

则会有 $0 \leqslant \widetilde{\zeta}_n \xrightarrow{\mathbf{P}} \zeta$ 和 $\mathbf{E}\zeta_n \to \mathbf{E}\zeta$. 所以由第 17 题中的断言 (a), 随机变量族 $\{\zeta_n,\ n \geqslant 1\}$ 一致可积. 又由于 $0 \leqslant \widetilde{\xi}_n \leqslant \widetilde{\zeta}_n$, 所以 $\{\widetilde{\xi}_n,\ n \geqslant 1\}$ 也一致可积. 由此即得 $\mathbf{E}\widetilde{\xi}_n \to \mathbf{E}\widetilde{\xi}$ (并且还有 $\mathbf{E}|\widetilde{\xi}_n - \widetilde{\xi}| \to 0$). 再由条件 $\mathbf{E}\eta_n \to \mathbf{E}\eta$ 推得 $\mathbf{E}\xi_n \to \mathbf{E}\xi$.

20. 证明, $L_* f \leqslant L^* f$, 而如果函数 f 有界, 并且测度 μ 有限, 则 $L_* f = L^* f$ (参阅第 11 小节注 2).

21. 证明, 对于有界函数 f, 数学期望 $\mathbf{E} f = L_* f$ (参阅第 11 小节注 2).

22. 证明第 11 小节注 2 中的最后一个结论.

23. 设 $F = F(x)$ 是随机变量 X 的分布函数. 证明

(a) $\mathbf{E}|X| < \infty \quad \Leftrightarrow \quad \int_{-\infty}^{0} F(x)dx < \infty$ 并且 $\int_0^\infty (1 - F(x))dx < \infty$.

(b) $\mathbf{E}X^+ < \infty \quad \Leftrightarrow \quad \int_a^\infty \ln \frac{1}{F(x)} dx < \infty$, 对某个 $a > 0$.

提示: (b) 证明如下关系式:

$$\mathbf{E}[XI(X > a)] \leqslant \int_a^\infty \ln \frac{1}{F(x)} dx \leqslant \frac{1}{F(a)} \mathbf{E}[XI(X > a)], \quad a > 0.$$

24. 证明, 如果 $p > 0$, 并且 $\lim\limits_{x\to\infty} x^p \mathbf{P}\{|\xi| > x\} = 0$, 则对任何 $0 < r < p$ 都有 $\mathbf{E}|\xi|^r < \infty$. 试举例说明, 对于 $r = p$, 有可能 $\mathbf{E}|\xi|^p = \infty$.

25. 试举例说明, 存在这样的密度函数 $f(x)$, 它不是偶函数, 但是所有的奇数阶矩全都为 0: $\int_{-\infty}^{\infty} x^k f(x)dx = 0$, $k = 1, 3, \cdots$.

26. 试举例说明, 存在这样的随机变量 ξ_n, $n \geqslant 1$, 对其有

$$\mathbf{E} \sum_{n=1}^\infty \xi_n \neq \sum_{n=1}^\infty \mathbf{E}\,\xi_n.$$

27. 设随机变量 X 具有性质: 对任何 $\alpha > 1$, 都有

$$\frac{\mathbf{P}\{|X| > \alpha n\}}{\mathbf{P}\{|X| > n\}} \to 0, \quad n \to \infty.$$

证明, X 的任意阶矩都存在.

提示: 利用公式

$$\mathbf{E}|X|^N = N \int_0^\infty x^{N-1} \mathbf{P}\{|X| > x\} dx, \quad N \geqslant 1.$$

28. 设 X 为随机变量, 分别以概率 p_k 取值 $k = 0, 1, 2, \cdots$. 函数

$$G(s) = \mathbf{E}s^X \left(= \sum_{k=0}^\infty p_k s^k \right), \quad |s| \leqslant 1$$

称为 X 的*母函数*. (关于母函数的详细讨论见第 3 节附录.) 试验证如下各公式:

(a) 如果 X 为泊松随机变量, 亦即有 $p_k = e^{-\lambda}\lambda^k/k!$, 其中 $\lambda > 0$, $k = 0, 1, 2, \cdots$, 那么

$$G(s) = \mathbf{E}s^X = e^{-\lambda(1-s)}, \quad |s| \leqslant 1.$$

(b) 如果 X 为具有几何分布的随机变量, 亦即有 $p_k = pq^k$, 其中 $0 < p < 1$, $q = 1 - p$, $k = 0, 1, 2, \cdots$, 那么

$$G(s) = \mathbf{E}s^X = \frac{p}{1 - sq}, \quad |s| \leqslant 1.$$

(c) 如果 $X_1, \cdots, X_n$ 为独立同分布的随机变量, 有 $\mathbf{P}\{X_1 = 1\} = p$, $\mathbf{P}\{X_1 = 0\} = q$ $(q = 1 - p)$, 则有

$$G(s) = (ps + q)^n \quad \left(= \sum_{k=0}^n [C_n^k p^k q^{n-k}] s^k \right),$$

这表明 $\mathbf{P}\{X_1 + \cdots + X_n = k\} = C_n^k p^k q^{n-k}$.

29. 设 X 为取值于集合 $\{0, 1, 2, \cdots\}$ 的随机变量, 而 $G(s) = \sum_{k=0}^\infty p_k s^k$, 其中 $p_k = \mathbf{P}\{X = k\}$, $k \geqslant 0$.

对正整数 r[①], 证明

(a) 如果 $\mathbf{E}X^r < \infty$, 则阶乘矩 $\mathbf{E}(X)_r \equiv \mathbf{E}X(X-1)\cdots(X-r+1) < \infty$, 并且 $\mathbf{E}(X)_r = \lim\limits_{s\to 1} G^{(r)}(s)\ (= G^{(r)}(1))$, 其中 $G^{(r)}(s)$ 是 $G(s)$ 的第 r 阶导数.

(b) 如果 $\mathbf{E}X^r = \infty$, 则 $\mathbf{E}(X)_r = \infty$, 并且 $\lim\limits_{s\to 1} G^{(r)}(s) = \infty$.

30. 设 X 为离散随机变量, 服从集合 $\{0, 1, \cdots, n\}$ 中的均匀分布, 意即 $\mathbf{P}\{X = k\} = \frac{1}{n+1}$, $k = 0, 1, \cdots, n$. 先证明 $G(s) = \frac{1}{n+1} \frac{1 - s^{n+1}}{1 - s}$, 再求出 $\mathbf{E}X$ 和 $\mathbf{E}X^2$, 然后推出众所周知的恒等式

$$\sum_{k=1}^n k = \frac{n(n+1)}{2}, \quad \sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}.$$

31. (第一章第 1 节第 13 题续) 设 $A_1, \cdots, A_n$ 为某些事件 (未必相互独立), $X_i = I_{A_i}$, $i = 1, \cdots, n$, 而 $\Sigma_n = I_{A_1} + \cdots + I_{A_n}$. 证明, 母函数 $G_{\Sigma_n}(s) = \mathbf{E}s^{\Sigma_n}$ 由下式给出:

$$G_{\Sigma_n}(s) = \sum_{m=0}^n S_m (s-1)^m,$$

① 原书为: $r \geqslant 1$ —— 译者注.

其中

$$S_m = \sum_{1\leqslant i_1<\cdots<i_m\leqslant n} \mathbf{P}(A_{i_1}\cap\cdots\cap A_{i_m}) \left(= \sum_{1\leqslant i_1<\cdots<i_m\leqslant n} \mathbf{P}\{X_{i_1}=1,\cdots,X_{i_m}=1\}\right)$$

(参阅第一章第 1 节第 12 题). 并由此推出, 事件 $B_m=\{\Sigma_n=m\}$ 的概率由下式给出:

$$\mathbf{P}(B_m)=\sum_{k=m}^{n}(-1)^{k-m}C_k^m S_k.$$

提示: 利用 $G_{\Sigma_n}(s)=\mathbf{E}\prod_{i=1}^{n}(1+X_i(s-1))$ 和

$$\prod_{i=1}^{n}(1+X_i(s-1))=1+\sum_{i=1}^{n}X_i(s-1)+\sum_{1\leqslant i_1<i_2\leqslant n}X_{i_1}X_{i_2}(s-1)^2+\cdots+\prod_{i=1}^{n}X_i(s-1)^n.$$

32. 与母函数一样, 在众多场合起着有益作用的另一个概念是矩母函数: $M(s)=\mathbf{E}e^{sX}$ (在此自然需要假设 $\mathbf{E}e^{sX}<\infty$). (需要指出的是: 如果 X 是非负随机变量, 而 $\lambda\geqslant 0$, 则对 $s=-\lambda$, 函数 $\widehat{F}(\lambda)=M(-\lambda)\,(=\mathbf{E}e^{-\lambda X})$ 并非别的, 而恰恰是以 $F=F(x)$ 为分布函数的随机变量 X 的拉普拉斯变换.)

(a) 证明, 如果存在某个 $a>0$, 使得矩母函数 $M(s)$ 对所有 $s\in[-a,a]$ 都有定义, 则对一切 $k=1,2,\cdots$, 函数 $M(s)$ 在 $s=0$ 处的 k 阶导数 $M^{(k)}(s)$ 都存在, 并且

$$M^{(k)}(0)=\mathbf{E}X^k.$$

(正是这一性质使得人们将 $M(s)$ 称为矩母函数.)

(b) 举例说明, 存在随机变量, 它对于任何 $s>0$, 都有 $M(s)=\infty$.

(c) 证明, 若 X 为参数为 $\lambda>0$ 的泊松随机变量, 则对一切 $s\in\mathbb{R}$, 都有 $M(s)=e^{-\lambda(1-e^s)}$.

(d) 试举出两个相互独立的随机变量 X 与 Y 的例子, 对于它们, 矩母函数 $M_{X+Y}(s)=\mathbf{E}e^{s(X+Y)}$ 是矩母函数 $M_X(s)=\mathbf{E}e^{sX}$ 与 $M_Y(s)=\mathbf{E}e^{sY}$ 的乘积.

33. 设 $0<r<\infty$, $X_n\in L^r$, $X_n\xrightarrow{\mathbf{P}}X$. 则如下三条件相互等价 (试比较第 17 题与第 19 题):

(i) 随机变量族 $\{|X_n|^r,\ n\geqslant 1\}$ 一致可积.

(ii) $X_n\xrightarrow{L^r}X$ (即 $\mathbf{E}|X_n-X|^r\to 0$).

(iii) $\mathbf{E}|X_n|^r\to\mathbf{E}|X|^r<\infty$.

34. (斯皮策 (Spitzer) 恒等式) 设 $X_1,X_2,\cdots$ 为独立同分布随机变量, 有 $\mathbf{E}|X_1|<\infty$. 令 $S_k=X_1+\cdots+X_k$, $M_k=\max(0,S_1,\cdots,S_k)$, $k\geqslant 1$. 证明, 对任何 $n\geqslant 1$, 都有

$$\mathbf{E}M_n=\sum_{k=1}^{n}\frac{1}{k}\mathbf{E}S_k^+, \qquad (*)$$

其中 $S_k^+=\max(0,S_k)$.

文献[101]中的提示: 利用关系式

$$M_n=I(S_n>0)M_n+I(S_n\leqslant 0)M_n,$$
$$\mathbf{E}[I(S_n>0)M_n]=\mathbf{E}[I(S_n>0)X_1]+\mathbf{E}[I(S_n>0)M_{n-1}],$$
$$\mathbf{E}[I(S_n>0)X_1]=\frac{1}{n}\mathbf{E}S_n^+,$$

即可依次得到

$$\mathbf{E}M_n=\frac{1}{n}\mathbf{E}S_n^+ +\mathbf{E}M_{n-1}=\frac{1}{n}\mathbf{E}S_n^+ +\left(\frac{1}{n-1}\mathbf{E}S_{n-1}^+ +\mathbf{E}M_{n-2}\right)=\cdots$$
$$=\sum_{k=2}^{n}\frac{1}{k}\mathbf{E}S_k^+ +\mathbf{E}M_1=\sum_{k=1}^{n}\frac{1}{k}\mathbf{E}S_k^+.$$

注: 关系式 $(*)$ 可以通过对更为广泛的斯皮策等式中的变量 t 求导来得到: 对一切 $0<u<1$, 有

$$\sum_{k=0}^{\infty}u^k\mathbf{E}e^{itM_k}=\exp\left\{\sum_{k=1}^{\infty}\frac{u^k}{k}\mathbf{E}e^{itS_k^+}\right\},$$

对于该关系式的证明比直接证明 $(*)$ 更难.

35. 设 $S_0=0$, $S_n=X_1+\cdots+X_n$, $n\geqslant 1$, 为简单对称随机游动 (参阅第八章第 8 节), 而 $\sigma=\min\{n>0:\ S_n\geqslant 0\}$. 证明

$$\mathbf{E}\min(\sigma,2m)=2\mathbf{E}|S_{2m}|=4m\mathbf{P}\{S_{2m}=0\},\quad m\geqslant 0.$$

36. (a) 设 X 为标准高斯随机变量 ($X\sim\mathcal{N}(0,1)$). 利用分部积分证明, $\mathbf{E}X^k=(k-1)\mathbf{E}X^{k-2}$, $k\geqslant 2$. 并由此导出公式 ($k\geqslant 1$)

$$\mathbf{E}X^{2k-1}=0,\quad \mathbf{E}X^{2k}=1\cdot 3\cdot\cdots\cdot(2k-3)(2k-1)\ \ (=(2k-1)!!).$$

(b) 证明, 如果随机变量 X 服从参数 $\beta=1$ 的伽玛分布 (见第 3 节表 3), 则有

$$\mathbf{E}X^k=\frac{\Gamma(k+\alpha)}{\Gamma(\alpha)},\quad k\geqslant 1.$$

特别地, $\mathbf{E}X=\alpha$, $\mathbf{E}X^2=\alpha(\alpha+1)$, 因而, $\mathbf{D}X=\alpha$.

试对 $\beta\neq 1$ 的情形, 找出相应的公式.

(c) 证明, 如果随机变量 X 服从贝塔分布 (见第 3 节表 3), 则有

$$\mathbf{E}X^k=\frac{B(r+k,s)}{B(r,s)},\quad k\geqslant 1.$$

37. 设函数

$$\xi(\omega_1, \omega_2) = e^{-\omega_1\omega_2} - 2e^{-2\omega_1\omega_2}, \quad \omega_1 \in \Omega_1 = [1, \infty), \quad \omega_2 \in \Omega_2 = (0, 1].$$

证明,

(a) 对每个 ω_2, 函数 $\xi(\omega_1, \omega_2)$ 关于 $\omega_1 \in \Omega_1$ (勒贝格) 可积,

(b) 对每个 ω_1, 函数 $\xi(\omega_1, \omega_2)$ 关于 $\omega_2 \in \Omega_2$ (勒贝格) 可积,

但是富比尼定理不成立.

38. 证明莱维(B. Levy)定理: 设随机变量 $\xi_1, \xi_2, \cdots$ 可积 (对一切 $n \geqslant 1$, 有 $\mathbf{E}|\xi_n| < \infty$), $\sup\limits_n \mathbf{E}\xi_n < \infty$, 并且 $\xi_n \uparrow \xi$; 则随机变量 ξ 亦可积, 并且 $\mathbf{E}\,\xi_n \uparrow \mathbf{E}\,\xi$ (试比较定理 1a).

39. 证明法图引理的另一表述形式: 如果 $0 \leqslant \xi_n \to \xi$, ($\mathbf{P}$-a.s.), 且 $\mathbf{E}\,\xi_n \leqslant A < \infty$, $n \geqslant 1$. 则随机变量 ξ 亦可积, 且 $\mathbf{E}\,\xi \leqslant A$.

40. (关于勒贝格可积性与黎曼可积性之间的关系) 设博雷尔函数 $f = f(x)$, $x \in \mathbb{R}$ 关于勒贝格测度 λ 可积 ($\int_{\mathbb{R}} |f(x)|\lambda(dx) < \infty$). 证明, 对于任何 $\varepsilon > 0$, 都可以找到

(a) 阶梯函数 $f_\varepsilon(x) = \sum_{i=1}^n f_i I_{A_i}(x)$, 使得 $\int_{\mathbb{R}} |f(x) - f_\varepsilon(x)|\lambda(dx) < \varepsilon$, 其中 A_i 均为有限区间;

(b) 支撑有限的连续函数 $g_\varepsilon(x)$, 使得 $\int_{\mathbb{R}} |f(x) - g_\varepsilon(x)|\lambda(dx) < \varepsilon$.

41. 证明, 如果 ξ 是可积的随机变量, 则

$$\mathbf{E}\,\xi = \int_0^\infty \mathbf{P}\{\xi > x\}dx - \int_{-\infty}^0 \mathbf{P}\{\xi < x\}dx.$$

再证明, 对任何 $a > 0$, 都有

$$\mathbf{E}\,[\xi I(\xi > a)] = \int_a^\infty \mathbf{P}\{\xi > x\}dx + a\mathbf{P}\{\xi > a\},$$

而若 $\xi \geqslant 0$, 则还有

$$\mathbf{E}\,[\xi I(\xi \leqslant a)] = \int_0^a \mathbf{P}\{x < \xi \leqslant a\}dx.$$

42. 设 ξ 与 η 都是可积的随机变量. 证明

$$\mathbf{E}\xi - \mathbf{E}\eta = \int_{-\infty}^\infty [\mathbf{P}\{\eta < x \leqslant \xi\} - \mathbf{P}\{\xi < x \leqslant \eta\}]dx.$$

43. 设 ξ 是非负随机变量 ($\xi \geqslant 0$), 其拉普拉斯变换为 $\widehat{F}(\lambda) = \mathbf{E}e^{-\lambda\xi}$, $\lambda \geqslant 0$.

(a) 证明, 对一切 $0 < r < 1$, 都有

$$\mathbf{E}\,\xi^r = \frac{r}{\Gamma(1-r)} \int_0^\infty \frac{1 - \widehat{F}(\lambda)}{\lambda^{r+1}} d\lambda.$$

提示: 利用关系式: 对于 $s \geqslant 0$, $0 < r < 1$, 有

$$\frac{1}{r}\Gamma(1-r)s^r = \int_0^\infty \frac{1 - e^{-s\lambda}}{\lambda^{r+1}} d\lambda.$$

(b) 证明, 对一切 $r > 0$, 都有

$$\mathbf{E}\,\xi^{-r} = \frac{1}{r\Gamma(r)} \int_0^\infty \widehat{F}(\lambda^r) d\lambda.$$

提示: 利用关系式: 对于 $s \geqslant 0$, $r > 0$, 有

$$s = \frac{r}{\Gamma(1/r)} \int_0^\infty \exp\{-(\lambda/s)^r\} d\lambda.$$

44. (a) 证明, 第 7 小节中的赫尔德 (Hölder) 不等式 (29) 中的等号成立, 当且仅当, $|\xi|^p$ 与 $|\eta|^q$ 线性相关, 亦即存在均不为 0 的常数 a 和 b, 使得 $a|\xi|^p = b|\eta|^q$ ($\mathbf{P}$-a.s.).

(b) 证明, $1 < p < \infty$ 时的闵可夫斯基不等式 (31) 中的等号成立, 当且仅当, 对某均不为 0 的常数 a 和 b, 有 $a\xi = b\eta$ ($\mathbf{P}$-a.s.).

(c) 证明, 柯西–布尼亚科夫斯基不等式 (24) 中的等号成立, 当且仅当, ξ 与 η 线性相关 ($\mathbf{P}$-a.s.), 亦即存在均不为 0 的常数 a 和 b, 使得 $a\xi = b\eta$ ($\mathbf{P}$-a.s.).

45. 设 X 为随机变量, 有 $\mathbf{P}\{a \leqslant X \leqslant b\} = 1$, 其中 $a, b \in \mathbb{R}$, $a < b$. 记 $m = \mathbf{E}X$, $\sigma^2 = \mathbf{D}X$. 证明, $\sigma^2 \leqslant (m-a)(b-m)$, 并且等号成立, 当且仅当 $\mathbf{P}\{X = a\} + \mathbf{P}\{X = b\} = 1$.

46. 设 X 为随机变量, 有 $\mathbf{E}|X| < \infty$. 证明

(a) 如果 $X > 0$ ($\mathbf{P}$-a.s.), 则

$$\mathbf{E}\frac{1}{X} \geqslant \frac{1}{\mathbf{E}X}, \quad \mathbf{E}\ln X \leqslant \ln \mathbf{E}X, \quad \mathbf{E}(X\ln X) \geqslant \mathbf{E}X \cdot \ln \mathbf{E}X$$

($0 \cdot \ln 0 = 0$).

(b) 如果随机变量 X 的值属于区间 $[a, b]$, 其中 $0 < a < b < \infty$, 则有

$$1 \leqslant \mathbf{E}X \cdot \mathbf{E}\frac{1}{X} \leqslant \frac{(a+b)^2}{4ab},$$

并说明何时等号成立.

(c) 如果 X 为正值随机变量, 且 $\mathbf{E}X^2 < \infty$, 则有如下估计式 (佩利–济格蒙德不等式) 成立: 对 $0 < \lambda < 1$

$$\mathbf{P}\{X > \lambda\mathbf{E}X\} \geqslant (1-\lambda)^2 \frac{(\mathbf{E}X)^2}{\mathbf{E}X^2}.$$

(d) 如果对给定的 $u>0$, 存在 $c\geqslant 0$, 使得 $\mathbf{P}\{X\leqslant u\}\leqslant c$, 则对任何 $r>0$, 都有

$$\mathbf{E}X^r\leqslant\frac{u^r}{1-\dfrac{(c\mathbf{E}X^{2r})^{1/2}}{\mathbf{E}X^r}}$$

(假定分母有意义并且非 0).

(e) 证明, 如果 X 为非负随机变量, 则有

$$\mathbf{P}\{X>0\}\geqslant\frac{(\mathbf{E}X)^2}{\mathbf{E}X^2}.$$

47. 设 ξ 为随机变量, 有 $\mathbf{E}\xi=m$ 和 $\mathbf{E}(\xi-m)^2=\sigma^2$. 证明坎泰利不等式

$$\max\Big(\mathbf{P}\{\xi-m>\varepsilon\},\ \mathbf{P}\{\xi-m<\varepsilon\}\Big)\leqslant\frac{\sigma^2}{\sigma^2+\varepsilon^2},\quad\varepsilon>0,$$

$$\mathbf{P}\{|\xi-m|\geqslant\varepsilon\}\leqslant\frac{2\sigma^2}{\sigma^2+\varepsilon^2},\quad\varepsilon>0.$$

48. 设 ξ 为随机变量, 有 $\mathbf{E}|\xi|<\infty$, 而 $g=g(x)$ 为 $\mathbb{R}$ 上的严格下凸的函数. 证明, $\mathbf{E}g(\xi)=g(\mathbf{E}\xi)$ 当且仅当 $\xi=\mathbf{E}\xi$ ($\mathbf{P}$-a.s.).

49. 设 ξ 为可积的随机变量 ($\mathbf{E}|\xi|<\infty$). 证明, 对任何 $\varepsilon>0$, 都能找到一个简单随机变量 ξ_ε, 使得 $\mathbf{E}|\xi-\xi_\varepsilon|\leqslant\varepsilon$. (试比较《概率》第一卷第二章第 4 节定理 1 中的断言.)

50. 考察方程

$$Z_t=B(t)+\int_0^t Z_{s-}dA(s),\quad t\geqslant 0$$

(试比较第 6 节中的方程 (74)), 其中 $A(t)$ 与 $B(t)$ 为 ($t\geqslant 0$ 时的) 右连续函数, (在 $t>0$ 时) 存在左极限, 并且 (局部) 有界变差, $A(0)=B(0)=0$, $\Delta A(t)>-1$, 其中 $\Delta A(t)=A(t)-A(t-)$, $t>0$, $\Delta A(0)=0$.

证明, 所述的方程在 (局部) 有界变差函数类中, 存在如下形式的唯一解 $\mathscr{E}_t(A,B)$, 其中 $t>0$:

$$\mathscr{E}_t(A,B)=\mathscr{E}_{t-}(A)\int_0^t\frac{1}{\mathscr{E}_{s-}(A)}dB(s).$$

51. 设 (在 $t\geqslant 0$ 时) 右连续, (在 $t>0$ 时) 存在左极限, 并且 (局部) 有界变差的函数 $V(t)$ 满足积分不等式

$$V(t)\leqslant K+\int_0^t V(s-)dA(s),\quad t\geqslant 0,$$

其中 $K\geqslant 0$, $A(t)$ 为右连续和具有左极限的非降函数, 并且 $A(0)=0$.

证明,

$$V(t)\leqslant K\mathscr{E}_t(A),\quad t\geqslant 0;$$

特别地, 如果 $A(t)=\int\limits_0^t a(s)ds$, $a(s)\geqslant 0$, 则函数 $V(t)$ 满足格朗沃尔–贝尔曼 (Gronwall-Bellman) 不等式

$$V(t)\leqslant K\exp\left\{\int_0^t a(s)ds\right\},\quad t\geqslant 0.$$

52. 在导出赫尔德不等式 (29) 时曾经用到如下的不等式:

$$ab\leqslant\frac{a^p}{p}+\frac{b^q}{q}$$

($a>0$, $b>0$, 而 $p>1$, $q>1$, 使得 $\frac{1}{p}+\frac{1}{q}=1$). 证明, 该不等式 (在 $h(x)=x^{p-1}$ 时) 是如下的*杨不等式*的特例:

$$ab\leqslant H(a)+\widetilde{H}(b),\quad a>0,\ b>0,$$

其中

$$H(x)=\int_0^x h(y)dy,\quad\widetilde{H}(x)=\int_0^x\widetilde{h}(y)dy,$$

其中 $h=h(y)$, $y\in\mathbb{R}_+$ 是严格上升的连续函数, 有 $h(0)=0$, $\lim\limits_{y\to\infty}h(y)=\infty$, 而 $\widetilde{h}=\widetilde{h}(y)$, $y\in\mathbb{R}_+$ 是 $h=h(y)$ 的反函数, 亦即

$$\widetilde{h}(y)=\inf\{t:\ h(t)>y\}.$$

(在所考察的场合下, 连续的严格上升函数 $h=h(y)$ 具有反函数 $\widetilde{h}(y)=h^{-1}(y)$.)

53. 设 X 为随机变量. 证明, 对一切 $a>0$, 都有如下的相互蕴涵关系:

$$\mathbf{E}|X|^a<\infty\quad\Leftrightarrow\quad\sum_{n=1}^{\infty}n^{a-1}\mathbf{P}\{|X|\geqslant n\}<\infty.$$

54. 设 ξ 为非负随机变量. 证明, 对一切 $r>1$, 都有 (试与第 43 题进行比较):

$$\int_0^\infty\frac{\mathbf{E}(\xi\wedge x^r)}{x^r}dx=\frac{r}{r-1}\mathbf{E}\xi^{1/r}.$$

特别地,

$$\int_0^\infty\frac{\mathbf{E}(\xi\wedge x^2)}{x^2}dx=2\mathbf{E}\sqrt{\xi}.$$

55. 设 ξ 为非负随机变量, 有 $\mathbf{E}\xi\geqslant 0$, $0<\mathbf{E}\xi^2<\infty$, 又设 $\varepsilon\in[0,1]$. 证明如下的不等式 (*与切比雪夫不等式相反的不等式*):

$$\mathbf{P}\{\xi>\varepsilon\mathbf{E}\xi\}\geqslant(1-\varepsilon)^2\frac{(\mathbf{E}\xi)^2}{\mathbf{E}\xi^2}.^{①}$$

① 原书为: $\mathbf{P}\{\xi>\varepsilon\mathbf{E}\xi\}\geqslant(1-\varepsilon)\frac{2(\mathbf{E}\xi)^2}{\mathbf{E}\xi^2}$ —— 译者注.

56. 设 $(\Omega, \mathscr{F})$ 为任一可测空间. 在该空间上定义集合函数 $\mu = \mu(B)$, $B \in \mathscr{F}$, 如下:

$$\mu(B) = \begin{cases} |B|, & \text{如果 } B \text{ 为有限集}, \\ \infty, & \text{其余情形}, \end{cases}$$

其中 $|B|$ 为集合 B 的势 (元素的个数). 证明, 所定义的集合函数 μ (在第 1 节定义 6 的意义下) 是一个测度. 该测度称为计数测度, 它是 σ-有限的, 当且仅当 Ω 为至多可数的集合.

57. (针对拉东–尼科迪姆 (Radon-Nikodym) 定理. I.) 设 λ, μ 和 ν 都是可测空间 $(\Omega, \mathscr{F})$ 中的 σ-有限测度. 假定拉东–尼科迪姆导数 $\frac{d\nu}{d\mu}$ 和 $\frac{d\mu}{d\lambda}$ 存在. 证明, 导数 $\frac{d\nu}{d\lambda}$ 亦存在, 并且

$$\frac{d\nu}{d\lambda} = \frac{d\nu}{d\mu} \cdot \frac{d\mu}{d\lambda} \quad (\lambda\text{-a.s.}).$$

58. (针对拉东–尼科迪姆定理. II.) 考察可测空间 $(\Omega, \mathscr{F}) = ([0,1], \mathscr{B}([0,1]))$ 中的两个测度: 一个是勒贝格测度 λ, 另一个是计数测度 μ (参阅第 56 题). 证明, $\mu \ll \lambda$, 但是拉东–尼科迪姆定理中关于密度 $\frac{d\mu}{d\lambda}$ 存在的断言不满足.

59. 设 λ 和 μ 是可测空间 $(\Omega, \mathscr{F})$ 中的两个 σ-有限测度, $f = \frac{d\lambda}{d\mu}$. 证明, 如果 $\mu\{\omega: f = 0\} = 0$, 则密度 $\frac{d\mu}{d\lambda}$ 存在, 并且可将其取为

$$\varphi = \begin{cases} 1/f, & \text{在集合 } \{f \neq 0\} \text{ 上}, \\ c, & \text{在集合 } \{f = 0\} \text{ 上}, \end{cases}$$

其中 c 是任意选取的非负实数.

60. 证明, 在 $[0, \infty)$ 上定义的函数

$$f(x) = \begin{cases} 1, & x = 0, \\ \dfrac{\sin x}{x}, & x > 0 \end{cases}$$

是黎曼可积的 (并且 $(\mathrm{R})\int_0^\infty f(x)dx = \frac{\pi}{2}$), 但是却不是勒贝格可积的.

61. 试举例说明, 存在 $[0,1]$ 上的勒贝格可积的有界函数 $f = f(x)$, 对其不存在黎曼可积的函数 $g = g(x)$, 使得它在 $[0,1]$ 上几乎处处与 $f = f(x)$ 相等.

62. 试举例说明, 存在 $\mathbb{R}^2$ 中的博雷尔函数 $f = f(x,y)$, 它分别对 $y \in \mathbb{R}$ 和 $x \in \mathbb{R}$ 勒贝格可积, 有

$$\int_{\mathbb{R}} f(x,y)\lambda(dx) = 0, \quad \int_{\mathbb{R}} f(x,y)\lambda(dy) = 0,$$

但是却对于 $(\mathbb{R}^2, \mathscr{B}(\mathbb{R}^2))$ 中的勒贝格测度不可积.

63. (针对富比尼定理. I.) 设 $\lambda = \lambda(dx)$ 为勒贝格测度. $\mu = \mu(dy)$ 为 $[0,1]$ 上的计数测度. 以 D 表示正方形 $[0,1]^2$ 中的对角线. 证明

$$\int_{[0,1]}\Big(\int_{[0,1]} I_D(x,y)\lambda(dx)\Big)\mu(dy) = 0, \quad \int_{[0,1]}\Big(\int_{[0,1]} I_D(x,y)\mu(dy)\Big)\lambda(dx) = 1.$$

(这两个关系式表明, 在富比尼定理 (定理 8) 中, 关于测度有限的假定对于定理结论的成立是一个本质性的条件.)

64. (针对富比尼定理. II.) 证明, 如果将富比尼定理中关于 μ_1 与 μ_2 是有限测度的假设换为是 σ-有限的测度之后, 其结论仍然成立. 再证明, 如果连 σ-有限的假设都不加, 则一般来说, 富比尼定理的结论不再成立 (试比较第 63 题).

65. (针对定理 10 中的断言 a.) 试举例说明, 存在黎曼可积的有界非博雷尔函数. (本题本质上是第 3 节第 29 题的另一种表述.)

66. 设 $f = f(x)$ 是定义在赋有勒贝格测度 $\lambda = \lambda(dx)$ 的 $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$ 上的博雷尔函数. 假设 $\int_{\mathbb{R}^n} |f(x)|\lambda(dx) < \infty$. 证明

$$\lim_{h \to 0} \int_{\mathbb{R}^n} |f(x+h) - f(x)|\lambda(dx) = 0.$$

提示: 利用第 40 题中的断言 (b).

67. 我们知道, 对于任何有限个相互独立的可积随机变量 $\xi_1, \cdots, \xi_n$, 都有

$$\mathbf{E}\prod_{k=1}^n \xi_k = \prod_{k=1}^n \mathbf{E}\xi_k$$

(参阅定理 6). 证明, 如果 $\xi_1, \xi_2, \cdots$ 是独立的可积随机变量序列, 则一般来说, 会有

$$\mathbf{E}\prod_{k=1}^\infty \xi_k \neq \prod_{k=1}^\infty \mathbf{E}\xi_k.$$

68. 设 ξ 为随机变量, 现知其数学期望 $\mathbf{E}\xi < 0$, 并且对某个 $\theta \neq 0$, 有 $\mathbf{E}e^{\theta\xi} = 1$. 证明, 此时必有 $\theta > 0$.

69. 设 $h = h(t,x)$ 是定义在 $[a,b] \times \mathbb{R}$ 上的函数, 其中 $a, b \in \mathbb{R}$ 并且 $a < b$.

(a) 假设

1) 对于每个 $x^\circ \in \mathbb{R}$, 函数 $h(t, x^\circ)$, $t \in [a,b]$, 都是连续函数.

2) 对于每个 $t^\circ \in [a,b]$, 函数 $h(t^\circ, x)$, $x \in \mathbb{R}$, 都是 $\mathscr{B}(\mathbb{R})$ 可测的 (即为博雷尔函数).

证明, 函数 $h = h(t,x)$, $t \in [a,b], x \in \mathbb{R}$, 是 $\mathscr{B}([a,b]) \times \mathscr{B}(\mathbb{R})$ 可测的.

(b) 现在假设 ξ 为定义在某个概率空间上的随机变量, 并且假设除了上述条件 1) 和 2) 成立之外, 还有如下的条件成立:

3) 随机变量族 $\{h(t,\xi),\ t\in[a,b]\}$ 一致可积, 证明,

(i) 数学期望 $\mathbf{E}\,h(t,\xi)$ 关于 $t\in[a,b]$ 为连续函数.

(ii) 如果记 $H(t,x)=\int_a^t h(s,x)ds$, 则对于一切 $t\in(a,b)$, 都有导数 $\frac{d}{dt}\mathbf{E}H(t,\xi)$ 存在, 且等于 $\mathbf{E}\,h(t,\xi)$, 意即

$$\frac{d}{dt}\mathbf{E}\int_a^t h(s,\xi)ds=\mathbf{E}\,h(t,\xi).$$

70. (针对引理 2)(a) 设 ξ 为随机变量, 有 $\mathbf{E}|\xi|<\infty$. 证明, 对任何 $\varepsilon>0$, 都存在 $\delta>0$, 使得只要 $A\in\mathscr{F}$, $\mathbf{P}(A)<\delta$, 就有 $\mathbf{E}(|\xi|I_A)<\varepsilon$. 并由此推出: 对于满足条件 $\mathbf{E}|\xi|<\infty$ 的随机变量 ξ, 对任何 $\varepsilon>0$, 都存在常数 $K=K(\varepsilon)$, 使得

$$\mathbf{E}(|\xi|;\ |\xi|>K)\equiv\mathbf{E}|\xi|I(|\xi|>K)<\varepsilon.$$

(b) 设 $\{\xi_n,\ n\geqslant1\}$ 为一致可积的随机变量族. 证明, 随机变量族 $\{\frac{1}{n}\sum_{k=1}^n\xi_k,\ n\geqslant1\}$ 也是一致可积的.

71. 证明, 如果下凸函数 $g=g(x)$ 不是在整个空间 $\mathbb{R}$ 中给出, 而是仅在某个开集 $G\subseteq\mathbb{R}$ 中给出, 那么只要对随机变量 ξ, 有 $\mathbf{P}\{\xi\in G\}=1$ 和 $\mathbf{E}|\xi|<\infty$, 则延森 (Jensen) 不等式 (25) 仍然成立. 证明, 在开集 G 中给出的下凸函数 $g=g(x)$ 是该集合上的连续函数. 再证明, 任何这样的函数都可以表示为

$$g(x)=\sup_n(a_nx+b_n),\quad x\in G,$$

其中 a_n 与 b_n 为某些常数.

72. 证明, 对 $a,b\in\mathbb{R}$ 和非负实数 r, 有如下的所谓 c_r 不等式成立:

$$|a+b|^r\leqslant c_r(|a|^r+|b|^r),$$

其中, 当 $r\leqslant1$ 时, $c_r=1$; 而当 $r>1$ 时, $c_r=2^{r-1}$.

73. 设 ξ 与 η 为非负随机变量, 对任何 $x>0$, 有

$$\mathbf{P}\{\xi\geqslant x\}\leqslant x^{-1}\mathbf{E}[\eta I(\xi\geqslant x)].$$

证明, 对一切 $p>1$, 都有

$$\mathbf{E}\xi^p\leqslant\left(\frac{p}{p-1}\right)^p\mathbf{E}\eta^p.$$

提示: 首先考虑有界随机变量 $\xi_c=\xi\wedge c,\ c>0$. 根据 (69) 式, 对其有

$$\mathbf{E}\xi_c^p=p\int_0^c x^{p-1}\mathbf{P}\{\xi>x\}dx.$$

利用题中条件, 可得关于 $\mathbf{E}\xi_c^p$ 的不等式, 再令 $c\uparrow\infty$, 即得所证.

74. 证明如下的积分换元公式 (参阅第 15 题和关于勒贝格积分换元的定理 7):

设 I 是 $\mathbb{R}^n$ 中的开集, $y=\varphi(x)$ 为定义在 I 上的取值于 $\mathbb{R}^n$ 的函数 (如果 $x=(x_1,\cdots,x_n)\in I$, 则 $y=(y_1,\cdots,y_n)$, 其中 $y_i=\varphi_i(x_1,\cdots,x_n)$, $i=1,\cdots,n$). 假设所有的偏导数 $\frac{\partial\varphi_i}{\partial x_j}$ 都存在并连续, 使得 $|J_\varphi(x)|>0$, $x\in I$, 其中 $J_\varphi(x)$ 是函数 φ 的雅可比行列式, 即

$$J_\varphi(x)=\det\left\|\frac{\partial\varphi_i}{\partial x_j}\right\|,\quad 1\leqslant i,j\leqslant n.$$

在上述假设之下, 集合 $\varphi(I)$ 是 $\mathbb{R}^n$ 中的开集, 函数 φ 有反函数 $h=\varphi^{-1}$, 此时, $J_h(y)$ 存在, 为 $\varphi(I)$ 中的连续函数, 且有 $|J_h(y)|>0$, $y\in\varphi(I)$.

证明, 对于一切非负可积函数 $g=g(x)$, $x\in I$, 都有下式成立:

$$\int_I g(x)dx=\int_{\varphi(I)}g(h(y))|J_h(y)|dy,$$

亦即

$$\int_I g(x)dx=\int_{\varphi(I)}g(\varphi^{-1}(y))|J_{\varphi^{-1}}(y)|dy$$

(将积分理解为关于 $\mathbb{R}^n$ 中的勒贝格测度的勒贝格积分).

75. 设 $F=F(x)$ 为分布函数, 有 $F(0)=0$, $F(1)=1$, 并且在区间 $[0,1]$ 上满足利普希茨条件: $|F(x)-F(y)|\leqslant L|x-y|$. 设 m 为 $[0,1]$ 上的测度, 有 $m(B)=\int_B dF(x)$, $B\in\mathscr{B}([0,1])$; 而 λ 为 $[0,1]$ 上的勒贝格测度. 证明

$$\frac{dm}{d\lambda}\leqslant L\quad(\lambda\text{-a.s.}).$$

76. 设 $g=g(x)$ 为定义在实数轴 $\mathbb{R}$ 上的某个区间 $[a,b]$ 上的函数. 假设其为下凸函数 (对 $x,y\in[a,b]$ 和任何 $0\leqslant\lambda\leqslant1$, 都有 $g((1-\lambda)x+\lambda y)\leqslant(1-\lambda)g(x)+\lambda g(y)$). 证明, 这样的函数在开区间 (a,b) 上连续. 由此可知, 这样的函数必为博雷尔函数.

提示: 由下凸性可知: 对任何 $x,y,z\in[a,b]$, $x<y<z$, 都有

$$\frac{g(y)-g(x)}{y-x}\leqslant\frac{g(z)-g(y)}{z-y},$$

由此不难推得 $g=g(x)$ 在开区间 (a,b) 中的连续性.

77. 设 $G(s)=\sum_{k=0}^\infty p_k s^k$ 为离散随机变量 X 的母函数, 其中 $p_k=\mathbf{P}\{X=k\}$, $k=0,1,2,\cdots$ (参阅第 28 题). 令

$$q_k=\mathbf{P}\{X>k\},\quad r_k=\mathbf{P}\{X\leqslant k\}\quad k=0,1,2,\cdots.$$

证明, 序列 $q=(q_k)_{k\geqslant 0}$ 和 $r=(r_k)_{k\geqslant 0}$ 的母函数分别由如下二式给出:

$$G_q(s)=\frac{1-G(s)}{1-s},\quad |s|<1,\qquad G_r(s)=\frac{G(s)}{1-s},\quad |s|<1.$$

78. (关于破产概率, 试比较第一章第 9 节) 设 $S_0=x,\ S_n=x+\xi_1+\cdots+\xi_n,\ n\geqslant 1$, 其中 $(\xi_k)_{k\geqslant 1}$ 为独立同分布随机变量序列, 有 $\mathbf{P}\{\xi_k=1\}=p,\ \mathbf{P}\{\xi_k=-1\}=q,\ p+q=1$. 假设 x 为整数, 并且 $0\leqslant x\leqslant A$. 令

$$\tau=\inf\{n\geqslant 0:\ S_n=0\ 或\ S_n=A\}$$

为 ("游戏") 随机游动的结束时刻 (关于两人游戏, 可参阅第一章第 9 节). 记 $p_x(n)=\mathbf{P}\{\tau=n,\ S_n=0\}$, 则 $p_x(n)$ 表示游戏在第 n 步结束并且 $S_n=0$ 的概率, 可将其解释为 "破产" 概率.

证明, 母函数 $G_x(s)=\sum\limits_{n=0}^{\infty}p_x(n)\,s^n$ 满足递推关系式

$$G_x(s)=psG_{x+1}(s)+qsG_{x-1}(s),$$

$G_0(s)=1,\ G_A(s)=0.$

并证明, 该递推关系式的解为

$$G_x(s)=\left(\frac{q}{p}\right)^x\frac{\lambda_1^{A-x}(s)-\lambda_2^{A-x}(s)}{\lambda_1^A(s)-\lambda_2^A(s)},$$

其中

$$\lambda_1(s)=\frac{1}{2ps}(1+\sqrt{1-4pqs^2}),\quad \lambda_2(s)=\frac{1}{2ps}(1-\sqrt{1-4pqs^2}).$$

79. 彩票的编号为十进制六位数, 自 000000 至 999999, 从中随机抽取一张彩票. 试求彩票号码的各位数字之和等于 21 的概率 P_{21}.

提示: 采用以考察母函数为基础的方法, 参阅附录第 3 节 (本书 p321). 本题的答案为 $P_{21}=0.04$.

80. 设随机变量 ξ 的密度函数 $f(x)$ 为单峰函数, 其最大值在点 x_0 处达到 (x_0 称为分布的众数或峰点), 并且在 x_0 左边非降, 在其右边非升.

证明如下的高斯不等式: 对一切 $\varepsilon>0$, 都有

$$\mathbf{P}\left\{|\xi-x_0|\geqslant\varepsilon\sqrt{\mathbf{E}\,|\xi-x_0|^2}\right\}\leqslant\frac{4}{9\varepsilon^2}.$$

提示: 如果 $g(y)$ 在 $y>0$ 处非升, 则对 $\varepsilon>0$, 有

$$\varepsilon^2\int_\varepsilon^\infty g(y)dy\leqslant\frac{4}{9}\int_0^\infty y^2g(y)dy.$$

利用这一不等式, 可以得知, 对 $\varepsilon>0$ 和 $d^2=\mathbf{E}\,|\xi-x_0|^2$, 有

$$\mathbf{P}\{|\xi-x_0|\geqslant\varepsilon d\}\leqslant\frac{4}{9}\frac{\mathbf{E}\,[|\xi-x_0|/d]^2}{\varepsilon^2}=\frac{4}{9\varepsilon^2}.$$

81. 设 $\xi_1,\cdots,\xi_n$ 为独立同分布随机变量序列, 有 $\mathbf{P}\{\xi_i>0\}=1$ 和 $\mathbf{D}(\ln\xi_1)=\sigma^2$. 证明, 对 $\varepsilon>0$, 有

$$\mathbf{P}\left\{\xi_1\cdots\xi_n\leqslant(\mathbf{E}\ln\xi_1)^n e^{n\varepsilon}\right\}\geqslant 1-\frac{\sigma^2}{n\varepsilon^2}.$$

提示: 利用切比雪夫不等式

$$\mathbf{P}\{|Y_n-\mathbf{E}\,Y_n|\leqslant n\varepsilon\}\geqslant 1-\frac{\mathbf{D}\,Y_n}{n^2\varepsilon^2},$$

在其中取 $Y_n=\sum\limits_{i=1}^n\ln\xi_i$.

82. 设 $\mathbf{P}$ 和 $\widetilde{\mathbf{P}}$ 是 $(\Omega,\mathscr{F})$ 中的两个概率测度, 其中 $\widetilde{\mathbf{P}}$ 关于 $\mathbf{P}$ 绝对连续 ($\widetilde{\mathbf{P}}\ll\mathbf{P}$), 密度为

$$\frac{d\widetilde{\mathbf{P}}}{d\mathbf{P}}\leqslant c\quad(\mathbf{P}\text{-a.s.}),$$

其中 c 为常数 (显然有 $c\geqslant 1$). 证明, 可以找到 $\alpha\in(0,1]$ 和概率测度 $\mathbf{Q}$, 使得

$$\mathbf{P}=\alpha\widetilde{\mathbf{P}}+(1-\alpha)\mathbf{Q}.$$

提示: 取某个 $C>c$, 并令 $\alpha=1/C$ 及

$$\mathbf{Q}(A)=\frac{1}{1-\alpha}\int_A\left(1-\alpha\frac{d\widetilde{\mathbf{P}}}{d\mathbf{P}}\right)d\mathbf{P}.$$

83. 设 ξ 与 η 为相互独立的随机变量, 并且 $\mathbf{E}\,\xi=0$. 证明, $\mathbf{E}\,|\xi-\eta|\geqslant\mathbf{E}\,|\eta|$.

84. 设 $\xi_1,\xi_2,\cdots$ 为取值于 $\mathbb{R}=(-\infty,\infty)$ 的独立同分布的随机变量. 令 $S_0=0,\ S_n=\xi_1+\cdots+\xi_n$. 以递归方式定义 "阶梯指数"(也称为 "阶梯时刻"):

$$T_0=0,\quad T_k=\inf\{n>T_{k-1}:\ S_n-S_{T_{k-1}}>0\},\quad k\geqslant 1,$$

如同惯常, 令 $\inf\ \varnothing=\infty$. 显然, 对一切 $n\geqslant 1$, 都有

$$\mathbf{P}\{T_1=n\}=\mathbf{P}\{S_1\leqslant 0,\ \cdots,\ S_{n-1}\leqslant 0,\ S_n>0\}.$$

证明, 随机变量 T_1 的母函数 $G(s)=\sum\limits_{n=1}^{\infty}f_ns^n$ (其中 $f_n=\mathbf{P}\{T_1=n\}$) 由下式给出:

$$G(s)=\exp\left\{-\sum_{n=1}^{\infty}\frac{s^n}{n}\mathbf{P}\{S_n>0\}\right\},\quad |s|<1.$$

85. 设条件与符号均同上题, 令

$$A=\sum_{n=1}^{\infty}\frac{1}{n}\mathbf{P}\{S_n\leqslant 0\},\quad B=\sum_{n=1}^{\infty}\frac{1}{n}\mathbf{P}\{S_n>0\},$$

证明

$$\mathbf{P}\{T_1 < \infty\} = \begin{cases} 1, & \text{如果 } B = \infty, \\ 1 - e^{-B}, & \text{如果 } B < \infty, \end{cases}$$

并且, 如果 $B = \infty$, 则

$$\mathbf{E}\,T_1 = \begin{cases} e^A, & \text{如果 } A < \infty, \\ \infty, & \text{如果 } A = \infty. \end{cases}$$

86. 如同第 84 题, $\xi_1, \xi_2, \cdots$ 为独立同分布的随机变量序列, 有 $\mathbf{E}\,\xi_1 > 0$. 记 $S_0 = 0$, $S_n = \xi_1 + \cdots + \xi_n$, 令

$$\tau = \inf\{n \geqslant 1: \ S_n > 0\}.$$

证明 $\mathbf{E}\,\tau < \infty$.

87. 设第 84 题中的条件成立, 而 $\mathscr{R}_N$ 为冲程, 即 $S_0, S_1, \cdots, S_N$ 所能取的不同值的个数. 令

$$\sigma(0) = \inf\{n > 0: \ S_n = 0\},$$

即第一次回归零点的时刻.

证明, 当 $N \to \infty$ 时, 有

$$\mathbf{E}\frac{\mathscr{R}_N}{N} \to \mathbf{P}\{\sigma(0) = \infty\}.$$

($\mathbf{P}\{\sigma(0) = \infty\}$ 是不回归零点的概率.) 试将本题结论与第一章第 9 节第 7 题的结论作比较.

注: 如果所考察的是简单随机游动 (即 $\mathbf{P}\{\xi_1 = 1\} = p$, $\mathbf{P}\{\xi_1 = -1\} = q$), 则根据第八章第 8 节第 16 题, 当 $N \to \infty$ 时, 有

$$\mathbf{E}\frac{\mathscr{R}_N}{N} \to |p - q|,$$

换言之

$$\mathbf{E}\frac{\mathscr{R}_N}{N} \to \begin{cases} p - q, & \text{如果 } p > 1/2, \\ 0, & \text{如果 } p = 1/2, \\ q - p, & \text{如果 } p < 1/2. \end{cases}$$

88. 设 $\xi_1, \xi_2, \cdots$ 为相互独立的区间 $(0, 1)$ 上均匀分布的随机变量序列. 令

$$\nu = \min\{n \geqslant 2: \ \xi_n > \xi_{n-1}\}, \quad \mu(x) = \min\{n \geqslant 1: \ \xi_1 + \cdots + \xi_n > x\},$$

其中 $0 < x \leqslant 1$. 证明,

(a) $\mathbf{P}\{\mu(x) > n\} = x^n/n!$, $n \geqslant 1$.

(b) $\operatorname{Law}(\nu) = \operatorname{Law}(\mu(1))$.

(c) $\mathbf{E}\,\nu = \mathbf{E}\,\mu(1) = e$.

提示: (a) 利用归纳法.

(b) 应当证明 $\mathbf{P}\{\nu > n\} = \mathbf{P}\{\xi_1 > \xi_2 > \cdots > \xi_n\} = 1/n!$.

89. 利用赫尔德不等式, 证明, 算术平均值不小于几何平均值: 对于 $x_i \geqslant 0$, $1 \leqslant i \leqslant n$, 有

$$\frac{1}{n}\sum_{i=1}^n x_i \geqslant \left(\prod_{i=1}^n x_i\right)^{1/n}$$

§7. 关于 σ-代数的条件概率和条件数学期望

1. 设 ξ 与 η 是独立同分布的随机变量, 数学期望 $\mathbf{E}\xi$ 确定. 证明

$$\mathbf{E}\{\xi \,|\, \xi + \eta\} = \mathbf{E}\{\eta \,|\, \xi + \eta\} = \frac{\xi + \eta}{2} \quad \text{(a.s.)}.$$

提示: 应当指出, 对任何 $A \in \sigma(\xi + \eta)$, 都有 $\mathbf{E}\xi I_A = \mathbf{E}\eta I_A$.

2. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布随机变量序列, 有 $\mathbf{E}|\xi_1| < \infty$. 证明

$$\mathbf{E}\{\xi_1 \,|\, S_n, S_{n+1}, \cdots\} = \frac{S_n}{n} \quad \text{(a.s.)},$$

其中 $S_n = \xi_1 + \cdots + \xi_n$.

提示: 利用事实: 对任何 $A \in \sigma(S_n, S_{n+1}, \cdots)$, 都有 $\mathbf{E}\xi_i I_A = \mathbf{E}\xi_j I_A$.

3. 设 X 与 Y 为随机变量, 使得正则分布 $P_x(B) = \mathbf{P}\{Y \in B \,|\, X = x\}$ 存在. 证明, 如果 $\mathbf{E}|g(X, Y)| < \infty$, 则 P_X-a.s. 地 (其中 $P_X(C) = \mathbf{P}\{\omega: \ X(\omega) \in C\}$) 有

$$\mathbf{E}[g(X, Y) \mid X = x] = \int g(x, y) P_x(dy).$$

提示: 利用正则条件测度的定义和 "π-λ 系" 的概念 (参阅第 2 节), 证明, 对任何形如 $\sum\limits_{i=1}^n \lambda_i I_{A_i}$ 的 $g(x, y)$, 映射

$$x \ \mapsto \ \int_{\mathbb{R}} g(x, y) P_x(dy)$$

都是 $\mathscr{B}(\mathbb{R})$ 可测的, 其中 $A_i \in \mathscr{B}(\mathbb{R}^2)$, 并且对于任何 $B \in \mathscr{B}(\mathbb{R})$, 都有

$$\mathbf{E}g(X, Y) I_B = \int_B \left(\int_{\mathbb{R}} g(x, y) P_x(dy)\right) Q(dx),$$

其中 Q 是随机变量 X 的分布. 由此首先可以推出, 对有界 $\mathscr{B}(\mathbb{R}^2)$ 可测的函数, 所证的性质成立; 然后可知它们对于任何使得 $\mathbf{E}|g(X, Y)| < \infty$ 成立的函数 $g(x, y)$ 都成立.

4. 设随机变量 ξ 的分布函数为 $F_\xi(x)$. 证明

$$\mathbf{E}(\xi \mid a<\xi\leqslant b)=\frac{\displaystyle\int_a^b x dF_\xi(x)}{F_\xi(b)-F_\xi(a)}$$

(假设 $F_\xi(b)-F_\xi(a)>0$).

提示: 根据条件期望的定义, 我们有

$$\mathbf{E}(\xi \mid a<\xi\leqslant b)=\frac{\mathbf{E}[\xi I(a<\xi\leqslant b)]}{\mathbf{E}[I(a<\xi\leqslant b)]}$$

(假设 $\mathbf{E}[I(a<\xi\leqslant b)]>0$).

5. 设 $g=g(x)$ 为定义在 $\mathbb{R}$ 上的下凸函数, 有 $\mathbf{E}|g(\xi)|<\infty$. 证明, 对于条件数学期望 ($\mathbf{P}$-a.s.) 成立延森不等式

$$g(\mathbf{E}(\xi\,|\,\mathscr{G}))\leqslant\mathbf{E}(g(\xi)\,|\,\mathscr{G}).$$

提示: 首先, 对于 ξ 关于 σ-代数 $\mathscr{G}$ 的正则条件分布 $Q(x;B)$, 有 (参阅定理 3)

$$\mathbf{E}(g(\xi)\,|\,\mathscr{G})(\omega)=\int_{\mathbb{R}} g(x)Q(\omega;dx),$$

然后运用对于 "正常" 数学期望的延森不等式.

6. 证明, 随机变量 ξ 与 σ-代数 $\mathscr{G}$ 相互独立 (意即对任何 $B\in\mathscr{G}$, 都有 ξ 与 $I_B(\omega)$ 独立), 当且仅当, 对于任何使得 $\mathbf{E}|g(\xi)|<\infty$ 成立的博雷尔函数 $g(x)$, 都有 $\mathbf{E}(g(\xi)\,|\,\mathscr{G})=\mathbf{E}g(\xi)$.

提示: 如果 $A\in\mathscr{G}$, 而 $B\in\mathscr{B}(\mathbb{R})$, 则由 ξ 与 $\mathscr{G}$ 相互独立, 可以推出 $\mathbf{P}(A\cap\{g(\xi)\in B\})=\mathbf{P}(A)\mathbf{P}\{g(\xi)\in B\}$, 这也就意味着 $\mathbf{E}(g(\xi)\,|\,\mathscr{G})=\mathbf{E}g(\xi)$. 反之, 若该等式成立, 则特别地, 可令 $g(\xi)=I(\xi\in B)$, 此时得到

$$\mathbf{P}(A\cap\{\xi\in B\})=\mathbf{P}(A)\mathbf{P}\{\xi\in B\},$$

再由 $A\in\mathscr{G}$ 与 $B\in\mathscr{B}(\mathbb{R})$ 的任意性, 即得 ξ 与 $\mathscr{G}$ 相互独立.

7. 设 ξ 为非负随机变量, $\mathscr{G}$ 为 σ-代数, $\mathscr{G}\subseteq\mathscr{F}$. 对于 $A\in\mathscr{G}$, 我们以等式 $\mathbf{Q}(A)=\int_A \xi d\mathbf{P}$ 定义测度 $\mathbf{Q}$. 证明, $\mathbf{E}(\xi\,|\,\mathscr{G})<\infty$ (a.s.), 当且仅当, 测度 $\mathbf{Q}$ 是 σ-有限的.

提示: 为证必要性, 需令 $A_n=\{\mathbf{E}(\xi\,|\,\mathscr{G})\leqslant n\}$, 验证 $\mathbf{Q}(A_n)\leqslant n$, 由此即可断言 $\mathbf{Q}$ 是 σ-有限的.

充分性. 由存在 $\mathscr{G}$ 中的集合 $A_1,A_2,\cdots$, 使得 $\bigcup_{i=1}^{\infty}A_i=\Omega$ 及 $\mathbf{Q}(A_i)<\infty$, $i=1,2,\cdots$ 的事实, 应当推出 $\mathbf{E}(\xi\,|\,\mathscr{G})<\infty$ ($\mathbf{P}$-a.s.).

8. 证明, 条件概率 $\mathbf{P}(A|B)$ 是 "连续的", 意即: 如果 $\lim_n A_n=A$, $\lim_n B_n=B$, $\mathbf{P}(B_n)>0$, $\mathbf{P}(B)>0$, 则有 $\lim_n\mathbf{P}(A_n|B_n)=\mathbf{P}(A|B)$.

9. 设 $\Omega=(0,1)$, $\mathscr{F}=\mathscr{B}((0,1))$, 而 $\mathbf{P}$ 为勒贝格测度. 设 $X(\omega)$ 与 $Y(\omega)$ 为相互独立的随机变量, 都服从区间 $(0,1)$ 上的均匀分布. 令 $Z(\omega)=|X(\omega)-Y(\omega)|$ 为 $X(\omega)$ 与 $Y(\omega)$ 之间的距离. 证明, Z 的分布 $F_Z(z)$ 具有密度 $f_Z(z)$, 并且 $f_Z(z)=2(1-z)$, $0\leqslant z\leqslant 1$. (由此可知, $\mathbf{E}Z=1/3$.)

10. 在半径为 R 的圆 $\{(x,y):x^2+y^2\leqslant R^2\}$ 中 "随机地" 选取两个点 A_1 和 A_2, 换言之, 我们相互独立地依密度 (以极坐标 $A_i=(\rho_i,\theta_i)$, $i=1,2$ 表示)

$$\mathbf{P}(\rho_i\in dr,\ \theta_i\in d\theta)=\frac{rdrd\theta}{\pi R^2},\quad i=1,2,$$

选取这两个点. 证明, A_1 与 A_2 之间的距离 ρ 具有分布密度 $f_\rho(r)$:

$$f_\rho(r)=\frac{2r}{\pi R^2}\left(2\arccos\left(\frac{r}{2R}\right)-\frac{r}{R}\sqrt{1-\left(\frac{r}{2R}\right)^2}\right),$$

其中 $0\leqslant r\leqslant 2R$.

11. 在 (以 $(0,0)$, $(0,1)$, $(1,1)$, $(1,0)$ 为顶点的) 单位正方形中 "随机地" (需明确含义!) 选取一点 $P=(x,y)$. 试求该点离点 $(1,1)$ 比离点 $(1/2,1/2)$ 更近的概率.

12. (会面问题) A, B 二人约定在 7 时至 8 时之间会面. 但二人都忘了会面的具体时间, 于是都 "随机地" 在 7 时至 8 时之间到达会面地点, 并且都等待了不多于 10 分钟. 证明, 他们能够见上面的概率不大于 11/36.

13. 设 $X_1,X_2,\cdots$ 为独立随机变量序列, $S_n=\sum_{i=1}^{n}X_i$. 证明, S_1 与 S_3 关于 S_2 所生成的 σ-代数 $\sigma(S_2)$ 条件独立.

14. 称 σ-代数 $\mathscr{G}_1$ 与 $\mathscr{G}_2$ 关于 σ-代数 $\mathscr{G}_3$ 条件独立, 如果

$$\mathbf{P}(A_1A_2\mid\mathscr{G}_3)=\mathbf{P}(A_1\mid\mathscr{G}_3)\mathbf{P}(A_2\mid\mathscr{G}_3),\quad 对一切\ A_i\in\mathscr{G}_i,\ \ i=1,2.$$

证明, $\mathscr{G}_1$ 与 $\mathscr{G}_2$ 关于 $\mathscr{G}_3$ 的条件独立性等价于下述条件中的任何一个 ($\mathbf{P}$-a.s.) 成立:

(a) $\mathbf{P}(A_1\mid\sigma(\mathscr{G}_2\cup\mathscr{G}_3))=\mathbf{P}(A_1\mid\mathscr{G}_3)$, 对一切 $A_1\in\mathscr{G}_1$.

(b) $\mathbf{P}(B\mid\sigma(\mathscr{G}_2\cup\mathscr{G}_3))=\mathbf{P}(B\mid\mathscr{G}_3)$, 对一切 $B\in\mathscr{P}_1$, 其中 $\mathscr{P}_1$ 为 π-系, 使得 $\mathscr{G}_1=\sigma(\mathscr{P}_1)$.

(c) $\mathbf{P}(B_1B_2\mid\sigma(\mathscr{G}_2\cup\mathscr{G}_3))=\mathbf{P}(B_1\mid\mathscr{G}_3)\mathbf{P}(B_2\mid\mathscr{G}_3)$, 对一切 $B_1\in\mathscr{P}_1$, $B_2\in\mathscr{P}_2$, 其中 $\mathscr{P}_1$ 和 $\mathscr{P}_2$ 为 π-系, 使得 $\mathscr{G}_1=\sigma(\mathscr{P}_1)$, $\mathscr{G}_2=\sigma(\mathscr{P}_2)$.

(d) $\mathbf{E}(X\mid\sigma(\mathscr{G}_2\cup\mathscr{G}_3))=\mathbf{E}(X\mid\mathscr{G}_3)$, 对一切 $\sigma(\mathscr{G}_2\cup\mathscr{G}_3)$ 可测的并且存在数学期望 $\mathbf{E}X$ 的随机变量 X (参阅第 6 节定义 2).

15. 证明, 关于条件数学期望的法图引理的如下推广形式 (试比较定理 2 中的 (d)):

设 $(\Omega,\mathscr{F},\mathbf{P})$ 为概率空间, $(\xi_n)_{n\geqslant1}$ 为随机变量序列, 使得数学期望 $\mathbf{E}\xi_n$, $n\geqslant1$ 和 $\mathbf{E}\,\underline{\lim}\,\xi_n$ 有定义 (可以取 $\pm\infty$ 值, 参阅第 6 节定义 2).

假设 $\mathscr{G}$ 是 $\mathscr{F}$ 中事件的子 σ-代数, 且

$$\sup_{n\geqslant 1}\mathbf{E}(\xi_n^- I(\xi_n\geqslant a)\mid\mathscr{G})\to 0\quad(\mathbf{P}\text{-a.s.}),\quad a\to\infty.$$

证明

$$\mathbf{E}(\underline{\lim}\,\xi_n\mid\mathscr{G})\leqslant\underline{\lim}\,\mathbf{E}(\xi_n\mid\mathscr{G})\quad(\mathbf{P}\text{-a.s.}).$$

16. 如同上题, $(\xi_n)_{n\geqslant 1}$ 为随机变量序列, 使得数学期望 $\mathbf{E}\xi_n$, $n\geqslant 1$ 有定义, $\mathscr{G}$ 是 $\mathscr{F}$ 中事件的子 σ-代数, 使得

$$\sup_n\lim_{k\to\infty}\mathbf{E}(|\xi_n|I(|\xi_n|\geqslant k)\mid\mathscr{G})=0\quad(\mathbf{P}\text{-a.s.}).\tag{$*$}$$

证明, 如果 $\xi_n\to\xi(\mathbf{P}\text{-a.s.})$, 并且 $\mathbf{E}\xi$ 有定义, 则

$$\mathbf{E}(\xi_n\mid\mathscr{G})\to\mathbf{E}(\xi\mid\mathscr{G})\quad(\mathbf{P}\text{-a.s.}).\tag{$**$}$$

17. 将上题中的条件 $(*)$ 换为: 对某个 $\alpha>1$, 有 $\sup\limits_{n\geqslant 1}\mathbf{E}(|\xi_n|^\alpha\mid\mathscr{G})<\infty\,(\mathbf{P}\text{-a.s.})$. 证明 $(**)$ 式中的收敛性仍然成立.

18. 设对某个 $p\geqslant 1$, 有 $\xi_n\xrightarrow{L^p}\xi$. 证明, 对任何子 σ-代数 $\mathscr{G}\subseteq\mathscr{F}$, 都有

$$\mathbf{E}(\xi_n\mid\mathscr{G})\xrightarrow{L^p}\mathbf{E}(\xi\mid\mathscr{G}).$$

19. 设 X 与 Y 为随机变量, 有 $\mathbf{E}X^2<\infty$, $\mathbf{E}|Y|<\infty$.

(a) 令 $\mathbf{D}(X\,|\,Y)\equiv\mathbf{E}[(X-\mathbf{E}(X\,|\,Y))^2\,|\,Y]$. 证明, $\mathbf{D}X=\mathbf{E}\mathbf{D}(X\,|\,Y)+\mathbf{D}\mathbf{E}(X\,|\,Y)$. (试比较第一章第 8 节第 22 题.)

(b) 证明, $\mathbf{cov}(X,\ Y)=\mathbf{cov}(X,\ \mathbf{E}(Y\,|\,X))$.

20. 试说明, 例 5 中的充分统计量 $T(\omega)=s(X_1(\omega))+\cdots+s(X_n(\omega))$ 是否为最小的.

21. 试证明因式分解表达式 (57) 的正确性.

22. 证明, 在第 10 小节的例 5 中, 有 $\mathbf{E}_\theta(X_i\,|\,T)=\frac{n+1}{2n}T$, 其中, 对于 $\omega=(x_1,\cdots,x_n)$, 有 $X_i(\omega)=x_i$, $i=1,\cdots,n$.

23. 设 A,B 与 $C_1,\cdots,C_n$ 为概率空间 $(\Omega,\mathscr{F},\mathbf{P})$ 中 σ-代数 $\mathscr{F}$ 中的事件. 设对任何 $i=1,\cdots,n$, 都有

$$\mathbf{P}(C_i)>0,\quad\mathbf{P}(A\,|\,C_i)\geqslant\mathbf{P}(B\,|\,C_i)$$

且 $\bigcup\limits_{i=1}^n C_i=\Omega$. 试问, 是否必有 $\mathbf{P}(A)\geqslant\mathbf{P}(B)$?

24. 设 X 与 Y 为随机变量, 有 $\mathbf{E}|X|<\infty$, $\mathbf{E}|Y|<\infty$, 以及 $\mathbf{E}(X\,|\,Y)\geqslant Y$ 和 $\mathbf{E}(Y\,|\,X)\geqslant X\,(\mathbf{P}\text{-a.s.})$. 证明, $X=Y(\mathbf{P}\text{-a.s.})$.

提示: 首先证明, 关于 "$\geqslant$" 场合下的证明可以归结为等号情形下的证明.

方法 1: 取函数 $g(u)=\arctan u$, 则有 $(X-Y)(g(X)-g(Y))\geqslant 0\,(\mathbf{P}\text{-a.s.})$, 同时容易验证 $\mathbf{E}[(X-Y)(g(X)-g(Y))]=0$. 因此有 $X=Y\,(\mathbf{P}\text{-a.s.})$.

方法 2: 分别验证 $\mathbf{E}\frac{X^++1}{Y^++1}=1$, $\mathbf{E}\frac{Y^++1}{X^++1}=1$, 再令 $Z=\frac{X^++1}{Y^++1}$, 并证明 $\mathbf{E}(\sqrt{Z}-\frac{1}{\sqrt{Z}})^2=1$. 由此可得 $\mathbf{P}\{X^+=Y^+\}=1$. 同理可证 $\mathbf{P}\{X^-=Y^-\}=1$.

注: 我们在此顺便指出, 不存在这样的随机变量 X 与 Y, 使得 $\mathbf{E}|X|<\infty$, $\mathbf{E}|Y|<\infty$, 并使得严格不等式 $\mathbf{E}(X\,|\,Y)>Y$ 和 $\mathbf{E}(Y\,|\,X)>X$ 成立. (因若存在这样的随机变量, 那么就会导致如下的矛盾: $\mathbf{E}X=\mathbf{E}\mathbf{E}(X\,|\,Y)>\mathbf{E}Y=\mathbf{E}\mathbf{E}(Y\,|\,X)>\mathbf{E}X$.)

25. 设 X 为随机变量, 服从几何分布:

$$\mathbf{P}\{X=k\}=pq^{k-1},\quad k=1,2,\cdots,\quad 0\leqslant p\leqslant 1,\quad q=1-p.\tag{$*$}$$

证明, 对于 $m,n\in\{1,2,\cdots\}$, 有

$$\mathbf{P}\{X>m+n\,|\,X>n\}=\mathbf{P}\{X>m\}.\tag{$**$}$$

(试给出该性质的一种解释.)

再证明其逆命题: 如果取值于集合 $\{1,2,\cdots\}$ 的离散随机变量满足性质 $(**)$, 那么它必服从几何分布 $(*)$.

(试比较第 45 题.)

26. 证明, 随机向量 (X,Y) 与 $(\widetilde{X},Y)$ 同分布 $(X,Y)\overset{\mathrm{d}}{=}(\widetilde{X},Y)$, 当且仅当, 对任何事件 A, 都有 $\mathbf{P}\{X\in A\,|\,Y\}=\mathbf{P}\{\widetilde{X}\in A\,|\,Y\}$ $(\mathbf{P}\text{-a.s.})$.

27. 设 X 与 Y 为相互独立的参数分别为 $\lambda>0$ 和 $\mu>0$ 的泊松随机变量. 证明, 条件分布 $\mathrm{Law}(X\,|\,X+Y)$ 为二项分布:

$$\mathbf{P}\{X=k\,|\,X+Y=n\}=C_n^k\left(\frac{\lambda}{\lambda+\mu}\right)^k\left(\frac{\mu}{\lambda+\mu}\right)^{n-k},\quad 0\leqslant k\leqslant n.$$

28. 设随机变量 ξ 服从区间 $[-a,b]$ 上的均匀分布, 其中 $a>0$, $b>0$. 令 $\mathscr{G}_1=\sigma(|\xi|)$, $\mathscr{G}_2=\sigma(\mathrm{sign}\xi)$. 试求条件概率 $\mathbf{P}(A\,|\,\mathscr{G}_i)$, $i=1,2$, 其中 $A=\{\xi>0\}$ 和 $A=\{\xi\leqslant\alpha\}$, 而 $\alpha\in[-a,b]$.

29. 试举例说明, 等式 $\mathbf{E}(\xi+\eta\,|\,\mathscr{G})=\mathbf{E}(\xi\,|\,\mathscr{G})+\mathbf{E}(\eta\,|\,\mathscr{G})$ $(\mathbf{P}\text{-a.s.})$ 并非恒能成立; 试比较第 4 小节中的性质 $\mathbf{D}^*$.

提示: 可能会出现这样的情况, $\mathbf{E}(\xi+\eta\,|\,\mathscr{G})$ 确定并且 $(\mathbf{P}\text{-a.s.})$ 为零, 而此时作为和的 $\mathbf{E}(\xi\,|\,\mathscr{G})+\mathbf{E}(\eta\,|\,\mathscr{G})$ 却是不确定的.

30. 在事件 $B\in\mathscr{F}$ 关于 σ-代数 $\mathscr{G}\subseteq\mathscr{F}$ 的条件概率 $\mathbf{P}(B\,|\,\mathscr{G})(\omega)$ 的定义中 (参阅第 2 小节中的定义 2) 并未假定 $\mathbf{P}(\cdot\,|\,\mathscr{G})(\omega)$ 概率为 1 地是 $(\Omega,\mathscr{F})$ 上的测度.

试举例说明, $\mathbf{P}(\cdot\,|\,\mathscr{G})(\omega)$ 的某些版本 (以正概率 $\mathbf{P}$) 不是测度.

31. 试举例说明, 存在相互独立的随机变量 X, Y 和 σ-代数 $\mathscr{G}$, 使得对某些集合 A 和 B, 在正测集上有

$$\mathbf{P}(X \in A,\ Y \in B \,|\, \mathscr{G})(\omega) \neq \mathbf{P}(X \in A \,|\, \mathscr{G})(\omega)\mathbf{P}(Y \in B \,|\, \mathscr{G})(\omega).$$

(换言之, 独立性不能自动保证条件独立性.)

32. 如果随机变量族 $\{\xi_n,\ n \geqslant 1\}$ 一致可积, 并且 $\xi_n \to \xi$ ($\mathbf{P}$-a.s.), 则 $\mathbf{E}\xi_n \to \mathbf{E}\xi$ (参阅第 6 节定理 4b)). 同时在 $|\xi_n| \leqslant \eta$, $\mathbf{E}\eta < \infty, n \geqslant 1$, 并且 $\xi_n \to \xi$ ($\mathbf{P}$-a.s.) 的条件之下, 还可证明 (定理 2a)) 条件数学期望的 ($\mathbf{P}$-a.s.) 收敛性: $\mathbf{E}(\xi_n \,|\, \mathscr{G}) \to \mathbf{E}(\xi \,|\, \mathscr{G})$, $n \to \infty$.

试举例说明, 如果将条件 "$|\xi_n| \leqslant \eta$, $\mathbf{E}\eta < \infty, n \geqslant 1$" 换为 "随机变量族 $\{\xi_n,\ n \geqslant 1\}$ 一致可积", 则未必就有 $\mathbf{P}(\xi_n \,|\, \mathscr{G}) \to \mathbf{P}(\xi \,|\, \mathscr{G})$ ($\mathbf{P}$-a.s.), $n \to \infty$. 类似地, 对于第 6 节定理 4 (即关于一致可积随机变量族的法图引理), 若将其中的性质 a) 换为条件数学期望, 则亦会有相应的情况出现. (亦可参阅上面第 15 ~ 17 题中所引述的命题.)

33. 取 $([0,1], \mathscr{F}, \lambda)$ 作为概率空间 $(\Omega, \mathscr{F}, \mathbf{P})$, 其中 λ 是勒贝格测度, $\mathscr{F}$ 为博雷尔集合类. 试举例说明, 存在子 σ-代数 $\mathscr{G} \subseteq \mathscr{F}$, 在它上面的条件数学期望 $\mathbf{E}(1 \,|\, \mathscr{G})(\omega)$ 存在形如狄利克雷函数的版本:

$$d(\omega) = \begin{cases} 1, & \omega \text{ 为无理数}, \\ 0, & \omega \text{ 为有理数}. \end{cases}$$

(事实上, 完全可能出现这样的情况, 即对于非常 "光滑" 的函数 $\xi = \xi(\omega)$, 例如 $\xi(\omega) \equiv 1$, 其条件数学期望 $\mathbf{E}(\xi \,|\, \mathscr{G})(\omega)$ 也可能会是 ω 的极不 "光滑" 的函数.)

34. 如果随机变量 ξ 的数学期望 $\mathbf{E}\xi$ 确定, 则根据性质 $\mathbf{G}^*$ (参阅第 4 小节), 应当有 $\mathbf{E}(\mathbf{E}(\xi \,|\, \eta)) = \mathbf{E}\xi$ (其中 η 为某个随机变量). 试举例说明, 存在这样的随机变量 ξ 和 η, 对于它们, 数学期望 $\mathbf{E}(\mathbf{E}(\xi \,|\, \eta))$ 确定, 但是 $\mathbf{E}\xi$ 不存在.

35. 设 $\Omega = \{0, 1, 2, \cdots\}$, 而 $\mathbf{P}_\theta\{\omega = k\} = \frac{e^{-\theta}\theta^k}{k!}$ $(k \geqslant 0)$ 是以 $\theta > 0$ 为参数的泊松分布. 证明, 对于参数 $1/\theta$, 不存在无偏估计 $T = T(\omega)$(意即 $\mathbf{E}_\theta|T| < \infty$, $\theta > 0$, 且对一切 $\theta > 0$, 都有 $\mathbf{E}_\theta T = 1/\theta$.)

36. 考虑概率-统计模型 $(\Omega, \mathscr{F}, \mathscr{P})$, 其中概率测度 $\mathbf{P}$ 的族 $\mathscr{P} = \{\mathbf{P}\}$ 是受控的. 设 $\mathscr{G} \subseteq \mathscr{F}$ 是某个充分的 σ-代数. 证明, 任何满足条件 $\mathscr{G} \subseteq \widehat{\mathscr{G}} \subseteq \mathscr{F}$ 的 σ-代数 $\widehat{\mathscr{G}}$ 都是充分的. (在伯克霍尔德 (Burkholder) 的专著 [18] 中, 有例子表明, 在不受控的场合下, 一般来说, 没有此结论.)

37. 证明, 下列各空间都是博雷尔的:

(a) $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$.

(b) $(\mathbb{R}^\infty, \mathscr{B}(\mathbb{R}^\infty))$.

(c) 波利希空间 (完备可分的距离空间).

38. 设 $(E, \mathscr{E})$ 是博雷尔空间. 证明, 存在可数生成的代数 $\mathscr{A}$, 使得 $\sigma(\mathscr{A}) = \mathscr{E}$.

39. (针对性质 $\mathbf{K}^*$) 设 η 是 $\mathscr{G}$ 可测的随机变量, ξ 是 $\mathscr{F}$ 可测的随机变量, 并且 $\mathbf{E}|\eta|^q < \infty$, $\mathbf{E}|\xi|^p < \infty$, 其中 $p > 1$, $\frac{1}{p} + \frac{1}{q} = 1$. 证明, $\mathbf{E}(\xi\eta \,|\, \mathscr{G}) = \eta\mathbf{E}(\xi \,|\, \mathscr{G})$.

40. 设 X 为随机变量, 其分布对称 (Law (X) =Law $(-X)$), 以 $F(x)$ 表示其分布函数. 试求条件分布 $\mathbf{P}(X \leqslant x \,|\, \sigma(|X|))(\omega)$, $x \in \mathbb{R}$, 其中 $\sigma(|X|)$ 是由随机变量 $|X|$ 生成的 σ-代数.

41. 设 A 与 B 为事件, 有 $\mathbf{P}(A) = \alpha$, $\mathbf{P}(B) = 1 - \beta$, 其中 $0 \leqslant \beta < 1$, $\beta \leqslant \alpha$. 证明

$$\frac{\alpha - \beta}{1 - \beta} \leqslant \mathbf{P}(A \,|\, B) \leqslant 1 - \beta.$$

42. 设 p_k 是一个家庭中有 k 个孩子的概率, 并且

$$p_0 = p_1 = a(< \frac{1}{2}), \quad p_k = (1 - 2a)2^{-(k-1)}, \quad k \geqslant 2.$$

(生男孩和女孩的概率都是 1/2.)

试在一个家庭已经有两个男孩的条件下, 求

(a) 该家庭仅有两个孩子的概率 $P_{(a)}$.

(b) 该家庭还有两个女孩的概率 $P_{(b)}$.

提示: 可以简单地利用贝叶斯公式解答本题. 为了检验, 我们指出 $P_{(a)} = 27/64$, $P_{(b)} = 81/512$.

43. 设 X 为随机变量, 其分布对称 $(X \overset{\mathrm{d}}{=} -X)$, 函数 $\varphi = \varphi(x)$, $x \in \mathbb{R}$ 使得 $\mathbf{E}|\varphi(X)| < \infty$. 证明

$$\mathbf{E}[\varphi(X) \,|\, |X|] = \frac{1}{2}[\varphi(|X|) + \varphi(-|X|)] \quad (\mathbf{P}\text{-a.s.}).$$

44. 设 X 为非负随机变量. 试求条件概率:

$$\mathbf{P}(X \leqslant x \,|\, \lfloor X \rfloor) \quad \text{和} \quad \mathbf{P}(X \leqslant x \,|\, \lceil X \rceil),$$

其中 $\lfloor X \rfloor$ 表示不超过 X 的最大整数 (在一般的文献中, 通常记为 $[X]$), 而 $\lceil X \rceil$ 表示不小于 X 的最小整数.

45. 如果随机变量 X 服从参数 $\lambda > 0$ 的指数分布, 即 $\mathbf{P}\{X > x\} = e^{-\lambda x}$, $x \geqslant 0$, 则对任何非负实数 x 与 y, 都有如下性质 (无记忆性) 成立:

$$\mathbf{P}(X > x + y \,|\, X > x) = \mathbf{P}(X > y).$$

证明, 对于广义非负 (即取值于 $[0, \infty]$ 的) 随机变量 X, 如果具有所述的性质, 则必有下述情况之一发生: 或者 $\mathbf{P}\{X = 0\} = 1$; 或者 $\mathbf{P}\{X = \infty\} = 1$; 或者服从参数为某个 $0 < \lambda < \infty$ 的指数分布.

提示: 令 $f(x) = \mathbf{P}\{X > x\}$, $x \geqslant 0$, 则由所述性质得 $f(x + y) = f(x)f(y)$. 于是所要证明的命题归结为证明, 该函数方程在右连续不超过 1 的非负函数范

围内只有如下三种形式的解: 或 $f(x) \equiv 0$; 或 $f(x) \equiv 1$; 或对某个 $0 < \lambda < \infty$, 有 $f(x) = e^{-\lambda x}$.

46. 设随机变量 X 与 Y 均具有有限二阶矩. 证明,

(a) $\mathbf{cov}(X, Y) = \mathbf{cov}(X, \mathbf{E}(Y|X))$.

(b) 如果 $\mathbf{E}(Y|X) = 1$, 则有

$$\mathbf{D}X \leqslant \mathbf{D}XY.$$

47. 设 $X_1, \cdots, X_n$ 为独立同分布随机变量. 在离散情况下, 以 $f(x;\theta)$ 表示其分布取值 x ($x \in \{0, 1, \cdots\}$) 的概率; 在绝对连续情况下, 以 $f(x;\theta)$ ($x \in \mathbb{R}$) 表示其分布密度, 其中 θ 为某个参数.

我们要在关于 $f(x;\theta)$ 的各种不同假设情况之下, 讨论关于参数 θ 的充分统计量 $T_n = T_n(X_1, \cdots, X_n)$ 的构造问题.

证明,

(a) 如果 $f(x;\theta) = \theta(1-\theta)^x$, $x \in \{0, 1, \cdots\}$ 和 $0 \leqslant \theta < 1$ (几何分布), 则关于 θ 的充分统计量是 $T_n = X_1 + \cdots + X_n$.

(b) 如果

$$f(x;\theta) = \frac{\Gamma(\theta+2)}{\Gamma(\theta)\Gamma(2)} x^{\theta-1}(1-x), \quad 0 \leqslant x \leqslant 1, \quad \theta > 0$$

(参数为 $\alpha = \theta$, $\beta = 2$ 的贝塔分布; 参阅第 3 节表 3), 则关于 θ 的充分统计量是 $T_n = X_1 \cdot \cdots \cdot X_n$.

(c) 如果

$$f(x;\theta) = \frac{1}{\theta} e^{-x/\theta}, \quad x \geqslant 0, \quad \theta > 0$$

(参数为 $\lambda = 1/\theta$ 的指数分布, 亦为参数为 $\alpha = 1$, $\beta = \theta$ 的伽玛分布; 参阅第 3 节表 3), 则关于 θ 的充分统计量是 $T_n = X_1 + \cdots + X_n$.

(d) 如果

$$f(x;\theta) = \frac{1}{\sqrt{2\pi}\theta} e^{-x^2/2\theta^2}, \quad x \in \mathbb{R}, \quad \theta > 0$$

(参数为 $\sigma = \theta$ 的正态分布 $\mathcal{N}(0, \sigma^2)$), 则关于 θ 的充分统计量是 $T_n = X_1^2 + \cdots + X_n^2$.

(e) 如果

$$f(x;\theta) = \frac{1}{\Gamma(\theta)\beta^\theta} x^{\theta-1} e^{-x/\beta}, \quad x \geqslant 0, \quad \theta > 0$$

(参数为 $\alpha = \theta$, $\beta > 0$ 的伽玛分布; 参阅第 3 节表 3), 则关于 θ 的充分统计量是 $T_n = X_1 + \cdots + X_n$.

设 $f(x;\alpha,\beta)$ 为贝塔分布 (参阅第 3 节表 3) 的密度函数, 且 $\alpha = \beta = \theta > 0$. 试求关于参数 θ 的充分统计量.

设

$$F(x;\theta) = \begin{cases} 1 - (\theta_1/x)^{\theta_2}, & x \leqslant \theta_1, \\ 0, & x < \theta_1 \end{cases}$$

为两参数 $\theta = (\theta_1, \theta_2)$ 的帕雷托 (Pareto) 分布, 其中 $\theta_1 > 0$, $\theta_2 > 0$ (参阅第三章第 6 节第 23 题中的定义). 试求关于参数 $\theta = (\theta_1, \theta_2)$ 的充分统计量.

§8. 随机变量 II

1. 试证明 (9), (10), (24), (27), (28), (34) $\sim$ (38) 式.

2. 设 $\xi_1, \cdots, \xi_n$ 为独立同分布随机变量, 具有分布函数 $F(x)$ (当密度函数存在时, 具有密度函数 $f(x)$). 令 $\overline{\xi} = \max\{\xi_1, \cdots, \xi_n\}$, $\underline{\xi} = \min\{\xi_1, \cdots, \xi_n\}$, $\rho = \overline{\xi} - \underline{\xi}$. 证明

$$F_{\overline{\xi},\underline{\xi}}(y, x) = \begin{cases} \Big(F(y)\Big)^n - \Big(F(y) - F(x)\Big)^n, & y > x, \\ \Big(F(y)\Big)^n, & y \leqslant x, \end{cases}$$

$$f_{\overline{\xi},\underline{\xi}}(y, x) = \begin{cases} n(n-1)\Big(F(y) - F(x)\Big)^{n-2} f(x) f(y), & y > x, \\ 0, & y \leqslant x, \end{cases}$$

$$F_\rho(x) = \begin{cases} n\displaystyle\int_{-\infty}^{\infty} [F(y) - F(y-x)]^{n-1} f(y)\,dy, & x \geqslant 0, \\ 0, & x < 0, \end{cases}$$

$$f_\rho(x) = \begin{cases} n(n-1)\displaystyle\int_{-\infty}^{\infty} [F(y) - F(y-x)]^{n-2} f(y-x) f(y)\,dy, & x \geqslant 0, \\ 0, & x < 0. \end{cases}$$

3. 设随机变量 ξ_1 与 ξ_2 相互独立, 分别服从参数为 $\lambda_1 > 0$ 和 $\lambda_2 > 0$ 的泊松分布. 证明

(a) $\xi_1 + \xi_2$ 服从参数为 $\lambda_1 + \lambda_2$ 的泊松分布.

(b) $\xi_1 - \xi_2$ 具有如下形式的分布:

$$\mathbf{P}\{\xi_1 - \xi_2 = k\} = e^{-(\lambda_1+\lambda_2)} \left(\frac{\lambda_1}{\lambda_2}\right)^{k/2} I_k(2\sqrt{\lambda_1\lambda_2}), \quad k = 0, \pm 1, \pm 2, \cdots,$$

其中

$$I_k(2x) = x^k \sum_{r=0}^{\infty} \frac{x^{2r}}{r!\Gamma(k+r+1)}$$

是 k 次第一类修正的贝塞尔函数.

提示: 证明 (b) 的一种方法是: 考察随机变量 $\xi_1 - \xi_2$ 的以级数形式表出的母函数 (详见附录第 3 节).

4. 在《概率》第一卷第 256 页中的 (4) 式中令 $m_1 - m_2 = 0$. 证明

$$f_{\frac{\xi}{\eta}}(z) = \frac{\sigma_1\sigma_2\sqrt{1-\rho^2}}{\pi(\sigma_2^2 z^2 - 2\rho\sigma_1\sigma_2 z + \sigma_1^2)}.$$

5. 设 ξ 与 η 为随机变量, 称 $\rho^*(\xi, \eta) = \sup_{u,v} \rho(u(\xi), v(\eta))$ 是它们的最大相关系数, 其中的上确界为对一切使得相关系数 $\rho(u(\xi), v(\eta))$ 确定的博雷尔函数 $u = u(x)$, $v = v(x)$ 所取. 证明, 随机变量 ξ 与 η 相互独立, 当且仅当 $\rho^*(\xi, \eta) = 0$. (亦可参阅下面的第 6 题.)

提示: 必要性显然. 为证充分性, 令 $u(\xi) = I_A(\xi)$, $v(\eta) = I_B(\eta)$, 其中, A 与 B 为任意的博雷尔集合. 由性质 $\sup_{u,v} \rho(u(\xi), v(\eta)) = 0$ 应当可以获知

$$\mathbf{P}\{\xi \in A,\ \eta \in B\} - \mathbf{P}\{\xi \in A\}\mathbf{P}\{\eta \in B\} = \rho(I_A(\xi),\ I_B(\eta)) = 0.$$

再由 A 与 B 的任意性即得 ξ 与 η 相互独立.

6. (第 5 题续) 设 $(\Omega, \mathscr{F}, \mathbf{P})$ 为概率空间, $\mathscr{F}_1$ 与 $\mathscr{F}_2$ 是 $\mathscr{F}$ 的两个子 σ-代数, 而 $L^2(\Omega, \mathscr{F}_i, \mathbf{P})$ 是 $\mathscr{F}_i$ 可测的二阶矩有限的随机变量的空间, $i = 1, 2$.

令

$$\rho^*(\mathscr{F}_1,\ \mathscr{F}_2) = \sup |\rho(\xi_1, \xi_2)|,$$

其中 $\rho(\xi_1, \xi_2)$ 是 ξ_1 与 ξ_2 的相关系数, 上确界为对一切 $\xi_i \in L^2(\Omega, \mathscr{F}_i, \mathbf{P})$, $i = 1, 2$ 的随机向量 (ξ_1, ξ_2) 所取.

(a) 证明, 如果 $\mathscr{F}_1 = \sigma(X_1)$ 与 $\mathscr{F}_2 = \sigma(X_2)$, 其中 X_1 与 X_2 是 $(\Omega, \mathscr{F}, \mathbf{P})$ 上的两个随机变量, 则

$$\rho^*(\sigma(X_1),\ \sigma(X_2)) = \sup |\rho(X_1,\ X_2)|.$$

(b) 设 $\mathscr{F}_1 = \bigvee_{i\in I} \mathscr{A}_i (= \bigcup_{i\in I} \mathscr{A}_i)$, $\mathscr{F}_2 = \bigvee_{i\in I} \mathscr{B}_i (= \bigcup_{i\in I} \mathscr{B}_i)$, 其中 I 是某个指标集, $\mathscr{A}_i$, $\mathscr{B}_i$ 是 σ-代数, 而所有的 σ-代数 $\sigma(\mathscr{A}_i,\ \mathscr{B}_i)$, $i \in I$, 为整体独立. (将 $\sigma(\mathscr{A}_i,\ \mathscr{B}_i)$ 理解为由所有集合 $A_i \in \mathscr{A}_i$ 和 $B_i \in \mathscr{B}_i$ 所生成的 σ-代数.) 证明

$$\rho^*\left(\bigvee_{i\in I} \mathscr{A}_i,\ \bigvee_{i\in I} \mathscr{B}_i\right) = \sup_{i\in I} \rho^*(\mathscr{A}_i,\ \mathscr{B}_i).$$

7. 设 $\mathscr{F}_1$ 与 $\mathscr{F}_2$ 是两个子 σ-代数. 我们引入下列各种混合特征指数 (其中的上确界为对所有的集合 $A \in \mathscr{F}_1$, $B \in \mathscr{F}_2$ 所取):

$$\alpha(\mathscr{F}_1,\ \mathscr{F}_2) = \sup |\mathbf{P}(A \cap B) - \mathbf{P}(A)\mathbf{P}(B)|,$$

$$\varphi(\mathscr{F}_1,\ \mathscr{F}_2) = \sup |\mathbf{P}(B\,|\,A) - \mathbf{P}(B)|, \quad \mathbf{P}(A) > 0,$$

$$\psi(\mathscr{F}_1,\ \mathscr{F}_2) = \sup \left|\frac{\mathbf{P}(A \cap B)}{\mathbf{P}(A)\mathbf{P}(B)} - 1\right|, \quad \mathbf{P}(A)\mathbf{P}(B) > 0,$$

再令

$$\beta(\mathscr{F}_1,\ \mathscr{F}_2) = \sup\ \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{M} |\mathbf{P}(A_i \cap B_j) - \mathbf{P}(A_i)\mathbf{P}(B_j)|,$$

其中的上确界为对所有这样的分割对 $\{A_1, \cdots, A_N\}$ 与 $\{B_1, \cdots, B_M\}$ 所取, 其中对一切 $1 \leqslant i \leqslant N$, $1 \leqslant j \leqslant M$, 有 $A_i \in \mathscr{F}_1$, $B_j \in \mathscr{F}_2$.

结合上题所定义的 $\rho^*(\mathscr{F}_1,\ \mathscr{F}_2)$, 证明下述各不等式:

$$\alpha(\mathscr{F}_1,\ \mathscr{F}_2) \leqslant \beta(\mathscr{F}_1,\ \mathscr{F}_2) \leqslant \varphi(\mathscr{F}_1,\ \mathscr{F}_2) \leqslant \psi(\mathscr{F}_1,\ \mathscr{F}_2)$$

和

$$\alpha(\mathscr{F}_1,\ \mathscr{F}_2) \leqslant \rho^*(\mathscr{F}_1,\ \mathscr{F}_2) \leqslant 2\varphi^{1/2}(\mathscr{F}_1,\ \mathscr{F}_2),$$

$$\rho^*(\mathscr{F}_1,\ \mathscr{F}_2) \leqslant 2\varphi^{1/2}(\mathscr{F}_1,\ \mathscr{F}_2)\varphi^{1/2}(\mathscr{F}_2,\ \mathscr{F}_1).$$

8. 设 $\tau_1, \cdots, \tau_k$ 为独立同分布的非负随机变量, 服从密度函数如下所示的指数分布:

$$f(t) = \lambda e^{-\lambda t}, \quad t \geqslant 0.$$

证明, 随机变量 $\tau_1 + \cdots + \tau_k$ 的分布具有如下的密度函数:

$$\frac{\lambda^k t^{k-1} e^{-\lambda t}}{(k-1)!}, \quad t \geqslant 0,$$

并且有

$$P(\tau_1 + \cdots + \tau_k > t) = \sum_{i=0}^{k-1} e^{-\lambda t}\frac{(\lambda t)^i}{i!}.$$

9. 设 $\xi \sim \mathcal{N}(0, \sigma^2)$. 证明, 对一切 $p \geqslant 1$, 都有

$$\mathbf{E}|\xi|^p = C_p \sigma^p,$$

其中

$$C_p = \frac{2^{p/2}}{\pi^{1/2}}\Gamma\left(\frac{p+1}{2}\right),$$

而 $\Gamma(s) = \int_0^\infty e^{-x}x^{s-1}dx$ 是欧拉–伽玛函数. 特别地 (参阅第 6 节第 36 题), 对于所有 $n \geqslant 1$, 都有

$$\mathbf{E}\xi^{2n} = (2n-1)!!\sigma^{2n}.$$

10. 设 ξ 与 η 为相互独立的随机变量, 使得 $\xi + \eta$ 与 ξ 同分布. 证明 $\eta = 0$ ($\mathbf{P}$-a.s.).

提示: 设 $\eta_1, \cdots, \eta_n$, $n \geqslant 1$ 相互独立, 均与 η 同分布, 并且都与 ξ 独立. 则在题中的条件下, 对任何 $n \geqslant 1$, 随机变量 $\xi + \eta_1 + \cdots + \eta_n$ 都与 ξ 同分布.

如果 ξ 与 η 的所有阶矩都存在, 则为证题中结论, 还可通过比较半不变量 $s_{\xi+\eta}^{(k)}$ 与 $s_{\xi}^{(k)}$, $k \geqslant 1$ 来实现.

11. 设随机变量对 (U,V) 服从单位圆 $\{(u,v):\ u^2+v^2\leqslant 1\}$ 中的均匀分布, 而 $W^2=U^2+V^2$. 令

$$X=U\sqrt{-\frac{2\ln W}{W}},\quad Y=V\sqrt{-\frac{2\ln W}{W}}.$$

证明, X 与 Y 是相互独立的正态 $\mathcal{N}(0,1)$ 随机变量.

12. 设 U 与 V 是相互独立的区间 $(0,1)$ 上均匀分布的随机变量. 定义

$$X=\sqrt{-\ln V}\cos(2\pi U),\quad Y=\sqrt{-\ln V}\sin(2\pi U).$$

证明, X 与 Y 相互独立, 且都服从正态 $\mathcal{N}(0,1)$ 分布.

13. 设 R 是正值随机变量, 服从瑞利分布, 其密度函数为

$$f_R(r)=\frac{r}{\sigma^2}\exp\left\{-\frac{r^2}{2\sigma^2}\right\},\quad r>0,$$

其中 $\sigma^2>0$; 而 θ 为服从 $(\alpha,\alpha+2\pi k)$ 上的均匀分布的随机变量, 其中 $k\in\mathbb{N}=\{1,2,\cdots\}$, $\alpha\in[0,2\pi)$.

证明, 随机变量 $X=R\cos\theta$ 与 $Y=R\sin\theta$ 相互独立, 并且都服从正态分布 $\mathcal{N}(0,\sigma^2)$.

14. 试举例说明, 存在随机变量 ξ 与 η, 它们都服从高斯分布, 但是它们的和 $\xi+\eta$ 不服从高斯分布.

15. 设 X 为随机变量, 服从柯西分布, 其密度函数为

$$f(x)=\frac{1}{\pi}\cdot\frac{a}{a^2+x^2}.$$

证明, 随机变量 $Y=X^n$ 的密度函数 $g(y)$ 为

$$g(y)=\begin{cases}\dfrac{ay^{-\frac{n-1}{n}}}{\pi n(a^2+y^{2/n})}, & n \text{ 为奇数},\ y\in\mathbb{R},\\[2mm] \dfrac{2ay^{-\frac{n-1}{n}}}{\pi n(a^2+y^{2/n})}, & n \text{ 为偶数},\ y\in\mathbb{R}.\end{cases}$$

16. 设 $F(x)$ 为分布函数. 证明, 对任何 $a>0$, 函数

$$G_1(x)=\frac{1}{a}\int_x^{x+a}F(u)du \quad \text{与} \quad G_2(x)=\frac{1}{2a}\int_{x-a}^{x+a}F(u)du$$

也都是分布函数.

17. 设随机变量 X 服从参数为 $\lambda>0$ 的指数分布 $(f_X(x)=\lambda e^{-\lambda x}I(x\geqslant 0))$.

(a) 试求 $Y=X^{1/\alpha}$, $\alpha>0$ 的分布密度 (相应的分布称为韦布尔 (Weibull) 分布).

(b) 试求 $Y=\ln X$ 的分布密度 (相应的分布称为二重指数分布, 亦可参阅第 48 题).

(c) 证明, 随机变量 X 的整数部分 $[X]$ 与分数部分 $\{X\}$ 相互独立, 并求出它们的分布.

18. 设随机变量 X 与 Y 具有形如 $f(x,y)=g(\sqrt{x^2+y^2})$ 的联合密度函数 $f(x,y)$.

(a) 试求随机变量 $\rho=\sqrt{X^2+Y^2}$ 与 $\theta=\arctan(Y/X)$ 的联合密度函数. 证明, 随机变量 ρ 与 θ 相互独立.

(b) 令 $U=(\cos\alpha)X+(\sin\alpha)Y$, $V=(-\sin\alpha)X+(\cos\alpha)Y$, $\alpha\in[0,2\pi]$. 证明, 随机变量 U 与 V 的联合密度函数就是 $f(x,y)$ (这一事实反映了向量 (X,Y) 在 "旋转变换" 之下的分布不变性).

19. 设 $X_1,\cdots,X_n$ 为独立同分布的随机变量, 具有连续分布函数 $F=F(x)$. 在我们的假定之下, 既然有 $\mathbf{P}\{X_i=X_j\}=0$, $i\neq j$ (参阅, 例如, 本节中的第 76 题等), 因此

$$\mathbf{P}\{\text{对某些 } i\neq j \text{ 有 } X_i=X_j\}=\mathbf{P}\Big(\bigcup_{i<j}\{X_i=X_j\}\Big)\leqslant\sum_{i<j}\mathbf{P}\{X_i=X_j\}=0.$$

这就表明, 概率为 1 地可将 $X_1(\omega),\cdots,X_n(\omega)$ 按大小 (唯一地) 严格排序.

如果将排序后所得的新随机变量记作 $X_1^{(n)},\cdots,X_n^{(n)}$(称为秩序统计量, 亦可参阅第三章第 13 节和第一章第 12 节第 8 题), 则概率为 1 地有

$$X_1^{(n)}(\omega)<\cdots<X_n^{(n)}(\omega)$$

和 (试比较第 2 题)

$$X_1^{(n)}(\omega)=\min\,(X_1(\omega),\cdots,X_n(\omega)),\cdots,X_n^{(n)}(\omega)=\max\,(X_1(\omega),\cdots,X_n(\omega)).$$

(a) 证明, 对于 $n\geqslant 2$ 和 $2\leqslant k\leqslant n$, 随机变量 $X_k^{(n)}$ 的分布函数 $F_k^{(n)}=F_k^{(n)}(x)$ 满足如下关系式:

$$F_k^{(n)}(x)=F(x)F_{k-1}^{(n-1)}(x)+(1-F(x))F_k^{(n-1)}(x).$$

以下补充假设分布函数 $F=F(x)$ 具有密度函数 $f=f(x)$.

证明,

(b) 随机变量 $X_k^{(n)}$ 的概率分布密度为:

$$nf(x)C_{n-1}^{k-1}[F(x)]^{k-1}[1-F(x)]^{n-k}.$$

(c) 随机变量 $X_1^{(n)},\cdots,X_n^{(n)}$ 的联合密度 $f^{(n)}(x_1,\cdots,x_n)$ 为

$$f^{(n)}(x_1,\cdots,x_n)=\begin{cases}n!f(x_1)\cdots f(x_n), & \text{如果 } x_1<\cdots<x_n,\\ 0, & \text{其他场合}.\end{cases}$$

(d) 如果 $f(x)=I_{[0,1]}(x)$(即随机变量 X_i 服从区间 $[0,1]$ 上的均匀分布), 则

$$\mathbf{E}X_r^{(n)}=\frac{r}{n+1},\quad \mathbf{cov}(X_r^{(n)},X_p^{(n)})=\frac{r(n-p+1)}{(n+1)^2(n+2)},\quad r\leqslant p.$$

20. 设 $\xi_1,\cdots,\xi_n$ 为独立同分布的随机变量, 服从正态分布 $\mathcal{N}(m,\sigma^2)$. 随机变量

$$\overline{\xi}=\frac{1}{n}\sum_{i=1}^n\xi_i \quad 与 \quad s_1^2=\frac{1}{n-1}\sum_{i=1}^n(\xi_i-\overline{\xi})^2,\quad n>1,$$

分别称为样本均值和样本方差. 证明

(a) $\mathbf{E}s_1^2=\sigma^2$.

(b) 样本均值 $\overline{\xi}$ 与样本方差 s_1^2 相互独立.

(c) $\overline{\xi}\sim\mathcal{N}(m,\sigma^2/n)$, 而 $(n-1)s_1^2/\sigma^2$ 服从自由度为 $n-1$ 的 χ^2 分布.

21. 设 $X_1,\cdots,X_n$ 为独立同分布的随机变量, ν 为与 $X_1,\cdots,X_n$ 独立的取值于集合 $\{1,\cdots,n\}$ 的随机变量. 令 $S_\nu=X_1+\cdots+X_\nu$, 证明

$$\mathbf{D}S_\nu=\mathbf{D}X_1\mathbf{E}\nu+(\mathbf{E}X_1)^2\mathbf{D}\nu,\quad \frac{\mathbf{D}S_\nu}{\mathbf{E}S_\nu}=\frac{\mathbf{D}X_1}{\mathbf{E}X_1}+\mathbf{E}X_1\frac{\mathbf{D}\nu}{\mathbf{E}\nu}.$$

22. 设 $M(s)=\mathbf{E}e^{sX}$ 是随机变量 X 的矩母函数 (参阅第 6 节第 32 题). 证明, 对任何 $s>0$, 都有 $\mathbf{P}\{X\geqslant 0\}\leqslant M(s)$.

23. 设 $X,X_1,\cdots,X_n$ 为独立同分布的随机变量, $S_n=\sum_{i=1}^n X_i$, $S_0=0$, $\overline{M}_n=\max_{0\leqslant j\leqslant n}S_j$, $\overline{M}=\sup_{n\geqslant 0}S_n$. 以 $\xi\overset{\mathrm{d}}{=}\eta$ 表示随机变量 ξ 与 η 同分布. 证明,

(a) $\overline{M}_n\overset{\mathrm{d}}{=}(\overline{M}_{n-1}+X)^+$, $n\geqslant 1$.

(b) 如果 $S_n\to\infty$ ($\mathbf{P}$-a.s.), 则 $\overline{M}\overset{\mathrm{d}}{=}(\overline{M}+X)^+$.

(c) 如果 $-\infty<\mathbf{E}X<0$, 且 $\mathbf{E}X^2<\infty$, 则

$$\mathbf{E}\overline{M}=\frac{DX-D(S+X)^-}{-2\mathbf{E}X}.$$

24. 在上题中的条件下, 对 $\varepsilon>0$, 令 $\overline{M}(\varepsilon)=\sup_{n\geqslant 0}(S_n-n\varepsilon)$. 证明 $\lim_{\varepsilon\downarrow 0}\varepsilon\overline{M}(\varepsilon)=(\mathbf{D}X)/2$.

25. 设 ξ 与 η 为相互独立的随机变量, 分别具有密度函数 $f_\xi(x)$, $x\in\mathbb{R}$ 和 $f_\eta(y)=I_{[0,1]}(y)$ (即随机变量 η 服从区间 $[0,1]$ 上的均匀分布). 证明, 此时 (36) 和 (37) 式分别具有下列形式:

$$f_{\xi\eta}(z)=\begin{cases}\displaystyle\int_z^\infty\frac{f_\xi(x)}{x}dx, & z\geqslant 0,\\ 0, & z<0,\end{cases}$$

$$f_{\xi/\eta}(z)=\begin{cases}\displaystyle\int_0^1 xf_\xi(zx)dx, & 0\leqslant z\leqslant 1,\\ \displaystyle\frac{1}{z^2}\int_0^1 xf_\xi(x)dx, & z>1,\\ 0, & z<0.\end{cases}$$

特别地, 如果随机变量 ξ 也服从区间 $[0,1]$ 上的均匀分布, 则

$$f_{\xi/\eta}(z)=\begin{cases}\dfrac12, & 0\leqslant z\leqslant 1,\\ \dfrac{1}{2z^2}, & z>1,\\ 0, & z<0.\end{cases}$$

26. (a) 设 ξ 与 η 为相互独立的随机变量, 均服从参数为 $\lambda>0$ 的指数分布. 证明, 随机变量 $\frac{\xi}{\xi+\eta}$ 服从区间 $[0,1]$ 上的均匀分布.

(b) 设 ξ 与 η 为相互独立的随机变量, 分别服从参数为 λ_1 和 λ_2 的指数分布, 其中 $\lambda_1\neq\lambda_2$. 证明, 随机变量 $\xi+\eta$ 的密度函数 $f_{\xi+\eta}(z)$ 为:

$$f_{\xi+\eta}(z)=\frac{e^{-z/\lambda_1}-e^{-z/\lambda_2}}{\lambda_1-\lambda_2}I_{(0,\infty)}(z).$$

27. 设 ξ 与 η 为相互独立的正态 $\mathcal{N}(0,1)$ 随机变量. 证明,

(a) 随机变量 ξ/η 与 $\xi/|\eta|$ 均服从密度函数为 $\frac{1}{\pi(1+x^2)}$, $x\in\mathbb{R}$ 的柯西分布.

(b) 随机变量 $|\xi|/|\eta|$ 具有密度函数 $\frac{2}{\pi(1+x^2)}$, $x\geqslant 0$.

28. 设 $X_1,\cdots,X_n$ 为相互独立的随机变量, 分别服从参数为 $\lambda_1,\cdots,\lambda_n$ 的指数分布, 其中 $\lambda_i\neq\lambda_j$, $i\neq j$. 令 $T_n=X_1+\cdots+X_n$. 证明, 概率 $\mathbf{P}\{T_n>t\}$ 可以表示为下述形式:

$$\mathbf{P}\{T_n>t\}=\sum_{i=1}^n a_{in}e^{-\lambda_i t}.$$

试求系数 a_{in}, $i=1,\cdots,n$.

29. 设 $\xi_1,\xi_2,\cdots$ 为随机变量序列, 有 $\mathbf{E}\xi_n=0$, $\mathbf{E}\xi_n^2=1$. 令 $S_n=\xi_1+\cdots+\xi_n$, $n\geqslant 1$. 证明, 存在正数 a 和 b, 使得

$$\mathbf{P}\{对某个\ n\geqslant 1,\ 有\ S_n\geqslant an+b\}\leqslant\frac{1}{1+ab}.$$

30. 设随机变量 ξ 只可取有限个非负实数值 $x_1, \cdots, x_k$. 证明,

$$\lim_{n\to\infty}(\mathbf{E}\xi^n)^{1/n} = \max(x_1, \cdots, x_k).$$

31. 设 ξ 与 η 都是以 $\{1, \cdots, n\}$ 为取值集合的相互独立的随机变量. 假设或有 $\mathbf{E}\xi < \infty$, 或有 $\mathbf{E}\eta < \infty$. 证明,

$$\mathbf{E}\min(\xi, \eta) = \sum_{k=1}^{\infty}\mathbf{P}\{\xi \geqslant k\}\mathbf{P}\{\eta \geqslant k\}.$$

32. 设 ξ_1 与 ξ_2 为相互独立的随机变量, 分别服从参数为 λ_1 与 λ_2 的指数分布. 试求随机变量 $\frac{\xi_1}{\xi_1+\xi_2}$ 与随机变量 $\frac{\xi_1+\xi_2}{\xi_1}$ 的分布函数.

33. 设 X 与 Y 为尺寸相同的随机矩阵, 并且 $\mathbf{E}YY^*$ 可逆. 证明如下的柯西–布尼亚科夫斯基矩阵不等式:

$$(\mathbf{E}XY^*)(\mathbf{E}YY^*)^{-1}(\mathbf{E}YX^*) \leqslant \mathbf{E}XX^*.$$

(符号 $\leqslant$ 表示该式左端与右端的矩阵之差为非负定阵.)

34. (舍朴 (L. Shepp)) 设 X 为伯努利随机变量, 有 $\mathbf{P}\{X=1\}=p$, $\mathbf{P}\{X=0\}=1-p$, $0<p<1$.

(a) 证明, 存在与 X 独立的随机变量 Y, 使得 $X+Y$ 的分布对称, 即 $X+Y \overset{\mathrm{d}}{=} -(X+Y)$.

(b) 在这样的随机变量中, 找出使得 $\mathbf{D}Y$ 达到最小的随机变量 Y.

35. 设 U 为随机变量, 服从区间 $(0,1)$ 上的均匀分布. 证明

(a) 对一切 $\lambda>0$, 随机变量 $-\frac{1}{\lambda}\ln U$ 服从参数为 λ 的指数分布.

(b) 随机变量 $\tan\pi\left(U-\frac{1}{2}\right)$ 服从以 $\frac{1}{\pi(1+x^2)}$, $x\in\mathbb{R}$ 为密度函数的柯西分布.

(c) 随机变量 $[nU]+1$ 服从集合 $\{1, 2, \cdots, n\}$ 上的离散均匀分布.

(d) 如果 $0<q<1$, 则随机变量 $X = 1+\left[\frac{\ln U}{\ln q}\right]$ 服从几何分布 $\mathbf{P}\{X=k\} = q^{k-1}(1-q)$, $k\geqslant 1$.

36. 试举例说明, 存在这样的独立同分布的非负随机变量序列 $X_1, X_2, \cdots$, 使得对任何 $p>0$, 都有 $\sup\limits_n \mathbf{E}X_n^p < \infty$, 但是对任何子列 $(n_1, n_2, \cdots)$, 却都有

$$\mathbf{P}\Big\{\sup_j X_{n_j} < \infty\Big\} = 0.$$

37. 设 ξ 与 η 为相互独立的随机变量, 分别具有分布函数 $F=F(x)$ 和 $G=G(x)$. 由性质 $\mathbf{P}\{\max(\xi,\eta)\leqslant x\} = \mathbf{P}\{\xi\leqslant x\}\mathbf{P}\{\eta\leqslant x\}$ 可以推知, 随机变量 $\max(\xi,\eta)$ 的分布函数为 $F(x)G(x)$. 试从另一途径重新建立这一结论, 亦即将事件 $\{\max(\xi,\eta)\leqslant x\}$ 表示为事件 $\{\xi\leqslant x, \ \xi\geqslant\eta\}$ 与事件 $\{\eta\leqslant x, \ \xi<\eta\}$ 的和, 并通过条件概率来计算它们的概率. (最后一步需要用到第 6 节中的 (68) 式.)

38. 设 ξ 与 η 为相互独立的随机变量, 使得它们的乘积 $\xi\eta$ 服从 (参数为 $\lambda>0$ 的) 泊松分布. 证明, ξ 与 η 之一必定在集合 $\{0,1\}$ 中取值.

39. 对于标准正态 $\mathcal{N}(0,1)$ 随机变量 ξ, 有

$$\mathbf{P}\{\xi\geqslant x\} \sim \frac{\varphi(x)}{x}, \quad x\to\infty, \quad 其中 \quad \varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}.$$

(参阅第一章第 6 节第 9 题.) 试对服从伽玛分布 (见第 3 节表 3) 的随机变量 γ, 找出相应的渐近性质.

40. 设 ξ 为随机变量, M_a 是它的中位数集合 (按照第一章第 4 节第 23 题 (a) 的定义; 那里所给出的关于离散随机变量的中位数的定义可以自动挪到任意随机变量上面). 证明, 对于任何 $b\in\mathbb{R}$ 和 $p\geqslant 1$, 只要 $\mathbf{E}|\xi|^p<\infty$, 就都有

$$|\mu-b|^p \leqslant 2\mathbf{E}|\xi-b|^p,$$

其中 $\mu=\mu(\xi)\in M_a$. (我们指出, 特别地, 如果 $\mathbf{E}|\xi|^2<\infty$, 则有 $|\mu-\mathbf{E}\xi|\leqslant\sqrt{2\mathbf{D}\xi}$.)

41. 设 ξ 与 η 为相互独立的随机变量, 都具有有限的二阶矩. 证明, 随机变量 $\xi+\eta$ 与 $\xi-\eta$ 不相关的充要条件是 $\mathbf{D}\xi=\mathbf{D}\eta$.

42. 设 f 与 g 是 L^1 中的两个函数. 用 $f*g$ 表示它们的卷积 $(=\int_{\mathbb{R}} f(y)g(x-y)dy)$. 证明如下的杨不等式:

$$\int_{\mathbb{R}}|(f*g)(x)|dx \leqslant \int_{\mathbb{R}}|f(x)|dx\int_{\mathbb{R}}|g(x)|dx.$$

43. (22) 式给出了用 ξ 的密度函数 $f_\xi(x)$ 来表示 $\eta=\varphi(\xi)$ 的密度函数 $f_\eta(y)$ 的公式:

$$f_\eta(y) = f_\xi(h(y))h'(y),$$

其中 $h(y)=\varphi^{-1}(y)$.

设 I 是 $\mathbb{R}^n$ 中的某个开集, 函数 $y=\varphi(x)$ 定义在 I 上取值于 $\mathbb{R}^n$. (即, 如果 $x=(x_1,\cdots,x_n)\in I$, 则 $y=(y_1,\cdots,y_n)$, 其中 $y_i=\varphi_i(x_1,\cdots,x_n)$, $i=1,\cdots,n$.) 假设所有的偏导数 $\frac{\partial\varphi_i}{\partial x_j}$ 都存在并且连续, 使得 $|J_\varphi(x)|>0$, $x\in I$, 其中 $J_\varphi(x)$ 是函数 φ 的雅可比行列式:

$$J_\varphi(x) = \det\left\|\frac{\partial\varphi_i}{\partial x_j}\right\|, \quad 1\leqslant i,j\leqslant n.$$

证明, 如果 $\xi=(\xi_1,\cdots,\xi_n)$ 是取值于 I 的随机向量, 具有密度函数 $f_\xi(x)$, 而 $\eta=\varphi(\xi)$, 则 η 的密度函数 $f_\eta(y)$ 存在, 并且在集合 $\varphi(I)=\{y:\ y=\varphi(x),\ x\in I\}$ 上, $f_\eta(y)$ 的值由如下公式给出:

$$f_\eta(y) = f_\xi(h(y))|J_h(y)|,$$

其中 $h=\varphi^{-1}$ 是 φ 的反函数 (并且, 由于 $|J_\varphi(x)|>0$, 故而 $|J_h(y)|>0$).

提示: 运用多元积分的换元公式 (第 6 节第 74 题), 在其中令 $g(x)=G(\varphi(x))f_\xi(x)$, 其中 G 是恰当选取的函数.

44. 设 $\eta=A\xi+B$, 其中 $\xi=(\xi_1,\cdots,\xi_n)$, $\eta=(\eta_1,\cdots,\eta_n)$, A 为 $n\times n$ 方阵, 有 $|\det A|>0$, 而 B 为 n 维向量. 证明

$$f_\eta(y)=\frac{1}{|\det A|}f_\xi(A^{-1}(y-B)).$$

提示: 令 $\varphi(x)=Ax+B$, 利用上面第 43 题的结论, 并需证明 $|J_{\varphi^{-1}}(y)|=1/|\det A|$.

45. (a) 设 ξ 与 η 为随机变量, 它们的相关系数是 $\rho(\xi,\eta)$. 证明

$$\rho(c_1\xi+c_2,\,c_3\eta+c_4)=\rho(\xi,\,\eta)\cdot\mathrm{sign}(c_1c_3),$$

其中

$$\mathrm{sign}(x)=\begin{cases}1, & x>0,\\ 0, & x=0,\\ -1, & x<0.\end{cases}$$

(b) 设 ξ_1,ξ_2,ξ_3,ξ_4 为随机变量, 相应的各相关系数分别为 $\rho(\xi_i,\,\xi_j)$, $i\neq j$. 证明

$$\rho(\xi_1+\xi_2,\,\xi_3+\xi_4)=[\rho(\xi_1,\,\xi_3)+\rho(\xi_1,\,\xi_4)]+[\rho(\xi_2,\,\xi_3)+\rho(\xi_2,\,\xi_4)].$$

46. 设 $X=(X_1,\cdots,X_n)$, $n\geqslant 2$ 为高斯向量, 它的各个分量是独立同分布的正态随机变量 $X_i\sim\mathcal{N}(0,\sigma^2), i=1,\cdots,n$. 我们来考察向量 $X=(X_1,\cdots,X_n)$ 的球坐标 $R,\Phi_1,\cdots,\Phi_{n-1}$, 亦即, 令 $R\geqslant 0$, $\Phi_i\in[0,2\pi)$, $1\leqslant i\leqslant n-1$, 和

$$\begin{aligned}&X_1=R\sin\Phi_1,\\ &X_m=R\sin\Phi_m\cos\Phi_{m-1}\cdots\cos\Phi_1,\quad 2\leqslant m<n-1,\\ &X_{n-1}=R\sin\Phi_{n-1}\cos\Phi_{n-2}\cdots\cos\Phi_1,\\ &X_n=R\cos\Phi_{n-1}\cos\Phi_{n-2}\cdots\cos\Phi_1.\end{aligned}$$

证明, 随机向量 $(R,\Phi_1,\cdots,\Phi_{n-1})$ 的密度函数 $f(r,\varphi_1,\cdots,\varphi_{n-1})$ 由如下的公式给出 (其中 $r\geqslant 0$, $\varphi_i\in[0,2\pi)$, $i=1,\cdots,n-1$, $n\geqslant 2$):

$$f(r,\varphi_1,\cdots,\varphi_{n-1})=\frac{r^{n-1}\exp\left(-\dfrac{r^2}{2\sigma^2}\right)}{(2\pi)^{n/2}\sigma^n}\cos^{n-2}\varphi_1\cos^{n-3}\varphi_2\cdots\cos\varphi_{n-2}.$$

47. 设 X 为取值于区间 $[0,1]$ 的随机变量, 它的分布函数为康托尔函数 (见第 3 节). 试求它的矩 $\mathbf{E}X^n$, $n\geqslant 1$.

48. (a) 证明, 下列各函数都是分布函数:

$$F_{\mathrm{G}}(x)=\exp(-e^{-x}),\qquad x\in\mathbb{R};$$

$$F_{\mathrm{F}}(x)=\begin{cases}0, & x<0,\\ \exp(-x^{-\alpha}), & x\geqslant 0,\end{cases}\quad \text{其中 } \alpha>0;$$

$$F_{\mathrm{W}}(x)=\begin{cases}\exp(-|x|^{-\alpha}), & x<0,\\ 1, & x\geqslant 0,\end{cases}\quad \text{其中 } \alpha>0,$$

其中, $F_{\mathrm{G}}(x)$ 被称为冈贝尔 (Gumbel) 分布 (或二重指数分布, 试比较第 17 题), $F_{\mathrm{F}}(x)$ 被称为弗雷歇 (Fréchet) 分布, $F_{\mathrm{W}}(x)$ 被称为韦布尔 (Weibull) 分布, 由它们所生成的如下三种类型的分布 F_{I}, F_{II}, F_{III}, 在极值理论中起着重要作用:

I 型分布 (冈贝尔型分布):

$$F_{\mathrm{I}}(x)=F_{\mathrm{G}}(ax+b),\quad a>0,\ b\in\mathbb{R};$$

II 型分布 (弗雷歇型分布):

$$F_{\mathrm{II}}(x)=F_{\mathrm{F}}(ax+b),\quad a>0,\ b\in\mathbb{R};$$

III 型分布(韦布尔型分布):

$$F_{\mathrm{III}}(x)=F_{\mathrm{W}}(ax+b),\quad a>0,\ b\in\mathbb{R}.$$

(b) 证明, 如果随机变量 X 具有 II 型分布, 则随机变量

$$Y=\ln(X-\mu)$$

具有 I 型分布. 相应地, 如果随机变量 X 具有 III 型分布, 则随机变量

$$Y=-\ln(\mu-X)$$

具有 I 型分布.

注: 这些性质告诉我们, I 型分布在 "极值理论" 中起着决定性的作用. 因此, I 型分布有时也被称为极值分布.

49. (阶乘矩) 设 X 为随机变量, 其阶乘矩 $m_{(r)}$ 的定义是:

$$m_{(r)}=\mathbf{E}X(X-1)\cdots(X-r+1),\quad r=1,2,\cdots,$$

亦即 $m_{(r)}=\mathbf{E}(X)_r$.

设随机变量 X 服从参数为 $\lambda>0$ 的泊松分布. 对于 $r=3$, 我们有 $m_{(3)}=\lambda$. 试对一切正整数 r, 求出 $m_{(r)}$.

50. 设随机变量 θ_1 与 θ_2 相互独立, 均服从区间 $[0,2\pi)$ 上的均匀分布. 令 $X_1 = \cos\theta_1$, $X_2 = \cos\theta_2$.

证明

$$\frac{1}{2}(X_1+X_2) \stackrel{\text{law}}{=} X_1X_2,$$

其中 $\stackrel{\text{law}}{=}$ 与 $\stackrel{\text{d}}{=}$ 的含义相同, 都表示 "同分布".

51. 设随机变量 θ 服从区间 $[0,2\pi)$ 上的均匀分布, 随机变量 C 服从密度函数为 $\frac{1}{\pi(1+x^2)}$, $x\in\mathbb{R}$ 的柯西分布.

(a) 证明, 随机变量 $\cos^2\theta$ 与随机变量 $\frac{1}{1+C^2}$ 同分布 (或者简记为 $\cos^2\theta \stackrel{\text{law}}{=} \frac{1}{1+C^2}$).

(b) 证明, $\cot\frac{\theta}{2} \stackrel{\text{law}}{=} C$.

(c) 试求随机变量 $\sin(\theta+\varphi)$ 与 $\alpha\tan\theta$ 的分布密度函数, 其中 $\varphi\in\mathbb{R}$, $\alpha>0$.

52. 设随机变量 ξ 服从标准指数分布, 意即 $\mathbf{P}\{\xi\geqslant t\} = e^{-t}$, $t\geqslant 0$, 而随机变量 N 服从标准正态分布, 即 $N\sim\mathcal{N}(0,1)$, 且 N 与 ξ 相互独立. 证明

$$\xi \stackrel{\text{law}}{=} \sqrt{2\xi}|N|,$$

意即随机变量 ξ 的分布 $\text{Law}\,(\xi)$ 与随机变量 $\sqrt{2\xi}|N|$ 的分布 $\text{Law}\,(\sqrt{2\xi}|N|)$ 相同.

53. 设 X 为随机变量, 所取之值为 $0,1,\cdots,N$, 它的二项矩为 $b_0,b_1,\cdots,b_N$ (其中 $b_k = C_X^k = \frac{1}{k!}\mathbf{E}(X)_k \equiv \frac{1}{k!}\mathbf{E}X(X-1)\cdots(X-k+1)$; 参阅附录中的第 3 节).

证明, X 的母函数为

$$G_X(s) = \mathbf{E}s^X = \sum_{k=0}^N b_k(s-1)^k = \sum_{n=0}^N s^n\left(\sum_{k=n}^N(-1)^{k-n}C_k^nb_k\right),$$

因而, 对每个 $n=0,1,\cdots,N$, 都有

$$\mathbf{P}\{X=n\} = \sum_{k=n}^N(-1)^{k-n}C_k^nb_k.$$

54. 设随机变量 X 与 Y 独立同分布, 都服从区间 $[0,1]$ 上的均匀分布. 证明, 随机变量

$$Z = \begin{cases} X+Y, & \text{如果 } 0\leqslant X+Y\leqslant 1,\\ (X+Y)-1, & \text{如果 } 1<X+Y\leqslant 2\end{cases}$$

也服从区间 $[0,1]$ 上的均匀分布.

55. 设随机变量 $X_1,\cdots,X_n$ 独立同分布, 都服从集合 $\{0,1,2\}$ 上的均匀分布 (即取其中每个值的概率都是 $1/3$). 令

$$P_n(k) = \mathbf{P}\{X_1+\cdots+X_n=k\}, \quad 0\leqslant k\leqslant 2n,$$

试求出 $P_n(k)$ 的一般表达式. (例如: $P_n(1) = n\cdot 3^{-n}$, $P_n(2) = C_{n+1}^2\cdot 3^{-n}$, $P_n(5) = (C_{n+4}^5 - nC_{n+1}^2)\cdot 3^{-n}$.)

56. 设 ξ 与 η 是任意两个随机变量, 有 $\mathbf{E}\xi^2<\infty$, $\mathbf{E}\eta^2<\infty$. 证明,

(a) $\mathbf{D}(\xi\pm\eta) = \mathbf{D}\xi+\mathbf{D}\eta\pm 2\mathbf{cov}(\xi,\eta)$.

(b) 如果 ξ 与 η 相互独立, 则有

$$\mathbf{D}(\xi\eta) = \mathbf{D}\xi\cdot\mathbf{D}\eta + \mathbf{D}\xi\cdot(\mathbf{E}\eta)^2 + \mathbf{D}\eta\cdot(\mathbf{E}\xi)^2.$$

(亦可参阅第 69 题中的命题.)

57. 设随机变量对 (X,Y) 的密度函数 $f(x,y)$ 是球对称的, 即对某个密度函数 $g=g(z)$, $z\geqslant 0$, 有

$$f(x,y) = g(x^2+y^2).$$

设 R 与 θ 是 (X,Y) 的极坐标, 即 $X=R\cos\theta$, $Y=R\sin\theta$. 证明, θ 服从区间 $[0,2\pi)$ 上的均匀分布, 而 R 具有密度函数 $h(r)=2\pi rg(r^2)$.

58. 设随机变量对 (X,Y) 具有密度函数. 定义复值随机变量

$$Z = X+iY, \quad Z_t = Ze^{it}, \quad t\in\mathbb{R}.$$

证明, 对任何 $t\in\mathbb{R}$, Z_t 服从同一个分布的必要条件是: $f(x,y)$ 具有 $f(x,y) = g(x^2+y^2)$ 的形式, 其中, 与上题一样, g 是某个密度函数.

59. 设随机变量 ξ 与 η 独立同分布, 具有指数分布密度函数 $f(x)=\lambda e^{-\lambda x}$, $x>0$. 证明, 随机变量 $\xi+\eta$ 与 ξ/η 相互独立.

60. 设随机变量 ξ 与 η 相互独立, 分别具有密度函数

$$f_\xi(x) = \frac{1}{\pi}\frac{1}{\sqrt{1-x^2}}, \quad |x|<1 \quad 与 \quad f_\eta(y) = \frac{y}{\sigma^2}e^{-\frac{y^2}{2\sigma^2}}, \quad y>0,\ \sigma>0.$$

证明, 随机变量 $\xi\eta$ 的分布是正态分布 $\mathcal{N}(0,\sigma^2)$.

61. 设 $\|\xi_{ij}\|$ 为 $n\times n$ 随机矩阵, 它的所有 (随机) 元素独立同分布, 均为 $\mathbf{P}\{\xi_{ij}=\pm 1\} = 1/2$. 证明, 该随机矩阵的行列式的均值与方差分别为 0 和 $n!$.

62. 设 $X_1,X_2,\cdots$ 为独立同分布的随机变量序列, 共同的分布为区间 $[0,1]$ 上的均匀分布. 证明,

$$\mathbf{E}\sum_{k=1}^n X_k^2\left(\sum_{k=1}^n X_k\right)^{-1} \to \frac{2}{3}, \quad n\to\infty.$$

63. 设随机向量 $X=(X_1,X_2,X_3)$ 服从如下的四面体中的均匀分布:

$$\Sigma_3 = \{(x_1,x_2,x_3):\ x_1\geqslant 0,\ x_2\geqslant 0,\ x_3\geqslant 0,\ x_1+x_2+x_3\leqslant c\},$$

其中 $c>0$. 试根据 $X=(X_1,X_2,X_3)$ 的分布, 分别求出随机变量 X_1 与随机向量 (X_1,X_2) 的分布. (这两个分布均称为 (X_1,X_2,X_3) 的边缘分布.)

提示: 首先证明 $X=(X_1,X_2,X_3)$ 的密度函数 $f(x_1,x_2,x_3)$ 是 V^{-1}, 其中 V 是 Σ_3 的体积, 其值为 $c^3/6$. 再由此出发, 推出二元 (边缘) 密度为 $f(x_1,x_2)=6(c-x_1-x_2)/c$, 最后再求出 (一元边缘) 密度 $f(x_1)$.

64. 设 $X_1,\cdots,X_n$ 为独立同分布的正值随机变量, 有 $\mathbf{E}X_1=\mu$, $\mathbf{E}X_1^{-1}=r$. 令 $S_m=X_1+\cdots+X_m$, $1\leqslant m\leqslant n$. 证明,

(a) $\mathbf{E}S_n^{-1}\leqslant r$.

(b) $\mathbf{E}X_iS_n^{-1}=1/n,\ i=1,\cdots,n$.

(c) $\mathbf{E}S_mS_n^{-1}=m/n$, 如果 $m\leqslant n$.

(d) $\mathbf{E}S_nS_m^{-1}=1+(n-m)\mathbf{E}S_m^{-1}$, 如果 $m<n$.

65. (狄利克雷分布) 在第 3 节表 3 中, (参数为 $\alpha>0,\ \beta>0$) 的贝塔分布被定义为区间 $[0,1]$ 上的具有如下的密度函数的分布:

$$f(x;\,\alpha,\,\beta)=\frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\,\beta)},$$

其中

$$B(\alpha,\,\beta)=\int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx\left(=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)},\ \text{而}\ \Gamma(\alpha)=\int_0^\infty x^{\alpha-1}e^{-x}dx\right).$$

多元贝塔分布就是狄利克雷分布, 它是集合

$$\Delta_{k-1}=\{(x_1,\cdots,x_{k-1}):\ x_i\geqslant 0,\ 0\leqslant x_1+\cdots+x_{k-1}\leqslant 1\},\quad k\geqslant 2$$

上的具有如下的密度函数的分布:

$$f(x_1,\cdots,x_{k-1};\,\alpha_1,\cdots,\alpha_{k-1},\alpha_k)=\frac{\Gamma(\alpha_1+\cdots+\alpha_k)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_k)}x_1^{\alpha_1-1}\cdots x_{k-1}^{\alpha_{k-1}-1}(1-(x_1+\cdots+x_{k-1}))^{\alpha_k-1},$$

其中 $\alpha_i>0,\ i=1,\cdots,k$. (有时也将该分布定义在单形 $\left\{(x_1,\cdots,x_k):\ x_i\geqslant 0,\ \sum\limits_{i=1}^k x_i=1\right\}$ 之上, 此时其 "密度函数" 为

$$\widetilde{f}(x_1,\cdots,x_{k-1};\,\alpha_1,\cdots,\alpha_{k-1},\alpha_k)=\frac{\Gamma(\alpha_1+\cdots+\alpha_k)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_k)}x_1^{\alpha_1}\cdots x_k^{\alpha_k}.$$

注意, 此处我们用了带引号的 "密度函数", 因为, 虽然函数 $\widetilde{f}(x_1,\cdots,x_{k-1};\,\alpha_1,\cdots,\alpha_{k-1},\alpha_k)$ 的形式唯一确定, 但是它却不是关于 $\mathbb{R}^n$ 中的勒贝格测度的密度.)

设 $X=(X_1,\cdots,X_k)$ 为随机向量, 有 $X_i\geqslant 0$ 和 $X_1+\cdots+X_k=1$, 服从狄利克雷分布, 在 Δ_{k-1} 上面具有密度函数 $f(x_1,\cdots,x_{k-1};\,\alpha_1,\cdots,\alpha_{k-1},\alpha_k)$ (这是 $X_1,\cdots,X_{k-1}$ 的密度函数, 其含义是: 将 $(X_1,\cdots,X_k)$ 中的最后一个随机变量 X_k 表示为 $X_k=1-(X_1+\cdots+X_{k-1})$, 从而只剩下 $k-1$ 个随机变量.)

(a) 证明

$$\mathbf{E}X_j=\frac{\alpha_j}{\sum\limits_{i=1}^k\alpha_i},\quad \mathbf{D}X_j=\frac{\alpha_j\left(\sum\limits_{i=1}^k\alpha_i-\alpha_j\right)}{\left(\sum\limits_{i=1}^k\alpha_i\right)^2\left(\sum\limits_{i=1}^k\alpha_i+1\right)},$$

$$\mathbf{cov}(X_{j_1},\,X_{j_2})=-\frac{\alpha_{j_1}\alpha_{j_2}}{\left(\sum\limits_{i=1}^k\alpha_i\right)^2\left(\sum\limits_{i=1}^k\alpha_i+1\right)},\quad j_1\neq j_2.$$

(b) 证明, 对非负数 $r_1,\cdots,r_k$, 我们有

$$\mathbf{E}X_1^{r_1}\cdots X_k^{r_k}=\frac{\Gamma\left(\sum\limits_{i=1}^k\alpha_i\right)\prod\limits_{i=1}^k\Gamma(\alpha_i+r_i)}{\prod\limits_{i=1}^k\Gamma(\alpha_i)\Gamma\left(\sum\limits_{i=1}^k(\alpha_i+r_i)\right)}.$$

(c) 试求随机变量 X_k 关于 $X_1,\cdots,X_{k-1}$ 的条件密度函数 $f_{X_k\,|\,X_1,\cdots,X_{k-1}}(x_k\,|\,x_1,\cdots,x_{k-1})$.

66. 设 X 为随机变量. 称如下的函数为 X 的浓度函数:

$$Q(X;\,l)=\sup_{x\in\mathbb{R}}\mathbf{P}\{x<X\leqslant x+l\},\quad l\geqslant 0.$$

(a) 证明, 如果随机变量 X 与 Y 相互独立, 则对一切 $l\geqslant 0$

$$Q(X+Y;\,l)\leqslant\min(Q(X;\,l),\ Q(Y;\,l)).$$

(b) 试找出 x_l^*, 使得 $Q(X;\,l)=\mathbf{P}\{x_l^*<X\leqslant x_l^*+l\}$, 并证明, X 的分布函数连续, 当且仅当 $Q(X;\,0)=0$.

提示: (a) 应当指出, 如果 X 与 Y 的分布函数分别为 F_X 与 F_Y, 则有

$$\mathbf{P}\{z<X+Y\leqslant z+l\}=\int_{-\infty}^{\infty}[F_X(z+l-y)-F_X(z-y)]dF_Y(y).$$

67. 设随机变量 $\xi\sim\mathcal{N}(m,\,\sigma^2)$(意即随机变量 ξ 服从参数为 m 与 σ^2 的正态分布). 令 $\eta=e^\xi$ 为对数正态随机变量, 则根据第 8 节中的 (23) 式, 其密度函数为

$$f_\eta(x)=\begin{cases}\dfrac{1}{\sqrt{2\pi}\,\sigma y}\exp\left[-\dfrac{(m-\ln y)^2}{2\sigma^2}\right], & y>0,\\ 0, & y\leqslant 0.\end{cases}$$

现对 $\alpha \in [-1,1]$, 构造函数

$$f^{(\alpha)}(y) = \begin{cases} f_\eta(y)\Big\{1-\alpha\sin[\pi\sigma^{-2}(m-\ln y)]\Big\}, & y>0, \\ 0, & y \leqslant 0. \end{cases}$$

(a) 证明, 函数 $f^{(\alpha)}(y)$ 是密度函数, 意即 $f^{(\alpha)}(y) \geqslant 0$ 且 $\int_0^\infty f^{(\alpha)}(y)dy = 1$.

(b) 设随机变量 ζ 具有密度函数 $f_\zeta(y) = f^{(\alpha)}(y)$, $\alpha \neq 0$. 证明, 随机变量 η 与 ζ 的所有各阶矩全都对应相等, 即 $\mathbf{E}\eta^n = \mathbf{E}\zeta^n$, $n \geqslant 1$. (事实上, 对数正态分布具有一切阶矩, 但是不能为这些矩所唯一确定.)

68. 设 $(\xi_n)_{n\geqslant 0}$ 为独立同分布随机变量序列, 其分布对称. 令 $S_0 = 0$, $S_n = \xi_1 + \cdots + \xi_n$, $n \geqslant 1$, 定义最大部分和序列 $M = (M_n)_{n\geqslant 0}$:

$$M_n = \max(S_0, S_1, \cdots, S_n),$$

与最小部分和序列 $m = (m_n)_{n\geqslant 0}$:

$$m_n = \min(S_0, S_1, \cdots, S_n).$$

证明 (此为第一章第 10 节第 7 题的推广): 对任何 n, 都有

$$(M_n - S_n,\ S_n - m_n,\ S_n) \overset{\text{law}}{=} (-m_n,\ M_n,\ S_n) \overset{\text{law}}{=} (M_n,\ -m_n,\ S_n),$$

即所述三个随机向量的分布相同.

提示: 利用如下的容易验证的性质: 对任何 n, 都有

$$(S_n - S_{n-k};\ k \leqslant n) \overset{\text{law}}{=} (S_k;\ k \leqslant n).$$

69. 设 ξ 与 η 为随机变量, 有 $\mathbf{D}\xi < \infty$ 和 $\mathbf{D}\eta < \infty$. 证明,

$$\mathbf{cov}^2(\xi,\ \eta) \leqslant \mathbf{D}\xi\mathbf{D}\eta,$$

并说明等号成立的条件.

70. 设 $\xi_1, \cdots, \xi_n$ 为独立同分布的随机变量. 证明

$$\mathbf{P}\{\min(\xi_1, \cdots, \xi_n) = \xi_1\} = n^{-1},$$

并证明, 随机变量 $\min(\xi_1, \cdots, \xi_n)$ 与 $I_{\{\xi_1 = \min(\xi_1, \cdots, \xi_n)\}}$ 相互独立.

71. 设随机变量 X 的分布函数是 $F = F(x)$, 而 C 为常数. 试求下列各随机变量的分布函数:

$$X \vee C \equiv \max(X,\ C), \quad X \wedge C \equiv \min(X,\ C), \quad X^C = \begin{cases} X, & \text{如果 } |X| \leqslant C, \\ 0, & \text{如果 } |X| > C. \end{cases}$$

72. 设 X 为随机变量, $\lambda > 0$, 而 $\varphi(x) = \frac{x}{1+x}$. 证明,

$$\mathbf{E}[\varphi(|X|^\lambda) - \varphi(x^\lambda)] \leqslant \mathbf{P}\{|X| \geqslant x\} \leqslant \frac{\mathbf{E}\varphi(|X|^\lambda)}{\varphi(x^\lambda)}.$$

73. 设随机变量 ξ 与 η 相互独立, 分别服从参数为 $(\alpha_1,\ \beta)$ 和 $(\alpha_2,\ \beta)$ 的伽玛分布 (参阅第 3 节表 3). 证明,

(a) 随机变量 $\xi + \eta$ 与 $\frac{\xi}{\xi+\eta}$ 相互独立.

(b) 随机变量 $\frac{\xi}{\xi+\eta}$ 服从参数为 (α_1, α_2) 的贝塔分布 (参阅第 3 节表 3).

74. (取胜概率随机的伯努利模型) 设 $\xi_1, \cdots, \xi_n$ 与 π 为随机变量, 其中 π 服从区间 $(0,1)$ 上的均匀分布, 而 ξ_i, $i = 1, \cdots, n$, 均以下述条件概率取 1 和 0 两个值:

$$\mathbf{P}\{\xi_i = 1 \mid \pi = p\} = p, \quad \mathbf{P}\{\xi_i = 0 \mid \pi = p\} = 1 - p,$$

并且条件独立, 即 (π-a.s.地) 有

$$\mathbf{P}\{\xi_1 = x_1, \cdots, \xi_n = x_n \mid \pi\} = \mathbf{P}\{\xi_1 = x_1 \mid \pi\} \cdots \mathbf{P}\{\xi_n = x_n \mid \pi\},$$

此处及下面, 对 $i = 1, \cdots, n$, 均有 $x_i = 0, 1$.

证明

(a) 无条件概率

$$\mathbf{P}\{\xi_1 = x_1, \cdots, \xi_n = x_n\} = \frac{1}{(n+1)C_n^x},$$

其中 $x = x_1 + \cdots + x_n$.

(b) 随机变量 $S_n = \xi_1 + \cdots + \xi_n$ 服从集合 $\{0, 1, \cdots, n\}$ 中的离散均匀分布.

(c) 条件分布 $\mathbf{P}\{\pi \leqslant p \mid \xi_1 = x_1, \cdots, \xi_n = x_n\}$ 与条件分布 $\mathbf{P}\{\pi \leqslant p \mid S_n = x_1 + \cdots + x_n\}$ 相同, $p \in (0,1)$.

(d) 分布 $\mathbf{P}\{\pi \leqslant p \mid S_n = x\}$, 其中 $x = x_1 + \cdots + x_n$, 具有密度

$$f_{\pi|S_n}(p \mid x) = (n+1)C_n^x p^x (1-p)^{n-x},$$

并且 $\mathbf{E}(\pi \mid S_n = x) = \frac{x+1}{n+2}$.

75. 设 ξ 与 η 为独立同分布的非负随机变量, 有 $\mathbf{P}\{\xi = 0\} < 1$. 假设 $\min(\xi, \eta)$ 与 $\xi/2$ 同分布. 证明, ξ 与 η 是指数分布随机变量.

提示: 由关系式

$$(\mathbf{P}\{\xi > x\})^2 = \mathbf{P}\{\min(\xi, \eta) > x\} = \mathbf{P}\{\xi > 2x\},$$

可以推出 $(\mathbf{P}\{\xi > x\})^{2n} = \mathbf{P}\{\xi > 2nx\}$. 由此可知, 对一切 $a > 0$ 和非负有理数 x, 都有 $\mathbf{P}\{\xi > x\} = e^{-\lambda x/a}$, 其中 $\lambda = -\ln \mathbf{P}\{\xi > a\}$. 接下来不难推知, 对任何 $x \geqslant 0$, 都有 $\mathbf{P}\{\xi > x\} = e^{-\lambda x/a}$.

76. 设 ξ 与 η 为独立同分布的随机变量, 具有分布函数 $F=F(x)$. 证明,

$$\mathbf{P}\{\xi=\eta\}=\sum_{x\in\mathbb{R}}|F(x)-F(x-)|^2.$$

(试比较第 12 节第 20 题.)

77. 设 $X_1,\cdots,X_n$ 为随机变量. 证明, 如下的 (关于随机变量最大值的) 容斥公式 (试比较第一章第 1 节第 12 题和第 4 节第 9 题):

$$\max(X_1,\cdots,X_n)=\sum_{i=1}^n X_i-\sum_{1\leqslant i_1<i_2\leqslant n}\min(X_{i_1},\ X_{i_2})$$
$$+\sum_{1\leqslant i_1<i_2<i_3\leqslant n}\min(X_{i_1},\ X_{i_2},\ X_{i_3})+\cdots+(-1)^{n+1}\min(X_1,\cdots,X_n).$$

并通过选择恰当的 $X_1,\cdots,X_n$, 由现在的容斥公式推出关于概率 $\mathbf{P}(A_1\cup\cdots\cup A_n)$ 的容斥公式. (参阅已经提到的第一章第 1 节第 12 题和第 4 节第 9 题.)

78. 设 $\xi_1,\xi_2,\cdots$ 为独立同分布的随机变量序列, 有 $\mathbf{P}\{\xi_1=1\}=\mathbf{P}\{\xi_1=-1\}=1/2$. 令 $z_n(\omega)=\frac{1}{2}\sum\limits_{k=1}^n\frac{\xi_k(\omega)}{3^k}$ 及 $z_\infty(\omega)=\lim\limits_{n\to\infty}z_n(\omega)$.

证明, 分布函数 $F(x)=\mathbf{P}\{z_\infty(\omega)\leqslant x\}$ 是康托尔函数 (参阅第 3 节). (事实上, 可以把 $\mathrm{Law}\,(z_\infty)$ 视为康托尔集上的康托尔测度.) 随机变量 $z_\infty=z_\infty(\omega)$ 是所谓奇异随机变量的例子 (其分布函数既不是离散的, 也不是绝对连续的, 见第 3 节第 18 题).

79. 证明, 在二项情形下 (第一章第 2 节第 1 小节及《概率》第一卷第 162 页中的表 2), 分布函数

$$B_n(m;\,p)\equiv\sum_{k=0}^m C_n^k p^k q^{n-k},\quad 0\leqslant m\leqslant n,$$

可以通过 (不完全) 贝塔函数表示为

$$B_n(m;\,p)=\frac{1}{\mathrm{B}(m+1,\ n-m)}\int_p^1 x^m(1-x)^{n-m-1}dx,$$

其中

$$\mathrm{B}(p,\,q)=\int_0^1 x^{p-1}(1-x)^{q-1}dx\quad\left(=\frac{\Gamma(p)\,\Gamma(q)}{\Gamma(p+q)}\quad\text{而}\quad\Gamma(p)=\int_0^\infty x^{p-1}e^{-x}dx\right).$$

80. 证明, 泊松分布函数 $F(m;\,\lambda)\equiv\sum\limits_{k=0}^m e^{-\lambda}\frac{\lambda^k}{k!}$, 其中 $m=0,1,2,\cdots$, 可以通过 (不完全) 伽玛函数表示为

$$F(m;\,\lambda)=\frac{1}{m!}\int_\lambda^\infty x^m e^{-x}dx.$$

81. 在利用均值与方差对分布的密度函数 $f=f(x)$ 进行描述时, 有两个标准化的特征参数, 一个是偏度系数 (skewness):

$$\alpha_3=\frac{\mu_3}{\sigma^3},$$

还有一个是峰度系数 (peakedness, kurtosis):

$$\alpha_4=\frac{\mu_4}{\sigma^4},$$

其中 $\mu_k=\int(x-\mu)^k f(x)dx,\ \mu=\int xf(x)dx,\ \sigma^2=\mu_2$.

试对表 3 中所列出的分布讨论它们的参数 α_3 和 α_4.

82. 设随机变量 X 服从参数为 n 和 p 的二项分布 (见《概率》第一卷第 162 页中的表 2). 证明, 对于它, 按照自然方式 (与第 81 题的定义方式类似) 所定义的 "偏度系数"

$$\mathrm{skw}(X)=\alpha_3\quad\left(=\frac{\mathbf{E}(X-\mathbf{E}X)^3}{(\mathbf{D}X)^{3/2}}\right)$$

满足下式 $(q=1-p)$:

$$\mathrm{skw}(X)=\frac{q-p}{\sqrt{npq}}.$$

(如果 $0<p<1/2$, 则 $\mathrm{skw}(X)>0$; 此时我们说, 分布具有右部 "长尾".) 再求该分布的峰度系数: $\mathrm{kur}(X)=\alpha_4\ \left(=\frac{\mathbf{E}(X-\mathbf{E}X)^4}{(\mathbf{D}X)^2}\right)$.

83. 设随机变量 $\xi_1,\cdots,\xi_n$ 独立同分布, 分别具有偏度系数 $\alpha_3\ (=\mathrm{skw}(\xi_1))$ 和峰度系数 $\alpha_4\ (=\mathrm{kur}(\xi_1))$. 证明

$$\mathrm{skw}(\xi_1+\cdots+\xi_n)=n^{-1/2}\mathrm{skw}(\xi_1)$$

和

$$\mathrm{kur}(\xi_1+\cdots+\xi_n)=3+n^{-1}\{\mathrm{kur}(\xi_1)-3\}.$$

84. 在引入二项分布的时候, 试验次数 n 是给定的, 人们考察的是概率 $\mathbf{P}_n\{\nu=r\}$, 亦即在 n 次试验中, 成功的次数 ν 等于 r 的概率. (该概率为 $\mathbf{P}_n\{\nu=r\}=C_n^r p^r q^{n-r},\ 0\leqslant r\leqslant n$, 其中, p 是每次试验成功的概率, 由此对所给定的 n 形成二项分布.) 负二项分布 (反二项分布) $\mathbf{P}^r\{\tau=k\}$(亦称为帕斯卡分布) 则是产生于这样的问题: "第 r 次成功首先出现在第 $\tau=k\,(\geqslant r)$ 次试验中的概率是多少?" 试证明,

$$\mathbf{P}^r\{\tau=k\}=C_{k-1}^{r-1}p^r q^{k-r},\quad k=r,r+1,\cdots,$$

其中 $r=1,2,\cdots$ (p 是每次试验中成功的概率). (根据定义, 对于给定的 r, 这些关于 $k=r,r+1,\cdots$ 的概率 $\mathbf{P}^r\{\tau=k\}$ 就形成了负二项分布.) 证明, (对于给定的 r) 有 $\mathbf{E}^r\tau=rq/p$.

85. 称取值于集合 $\{1,2,\cdots\}$ 的随机变量 ξ 服从参数为 $p>0$ 的 (离散的) 帕雷托 (Pareto) 分布, 如果

$$\mathbf{P}\{\xi=k\}=\frac{c}{k^{p+1}}.$$

证明

$$c=\frac{1}{\zeta(p+1)},\quad \mathbf{E}\xi=\frac{\zeta(p)}{\zeta(p+1)},$$

其中 $\zeta(s)=\sum\limits_{n=1}^{\infty}\frac{1}{n^s}$ 是黎曼 ζ 函数. (亦可参阅第三章第 6 节第 23 题关于连续的帕雷托分布的介绍.)

86. 设 $\xi_1,\cdots,\xi_n$ 为独立同分布的随机变量, 具有分布函数 $F(x\,|\,\theta)$. 该分布依赖于某个 (随机) 参数 θ, 而 θ 具有某个类 $\mathscr{K}$ 中的先验分布 $\Pi(\theta)$. 设 $\Pi(\theta\,|\,x_1,\cdots,x_n)$ 是根据 $\xi_1,\cdots,\xi_n$ 的观察值 $x_1,\cdots,x_n$ 按照贝叶斯公式计算出的后验分布. 如果后验分布仍然属于类 $\mathscr{K}$, 我们就说, $\mathscr{K}$ 中的分布 $\Pi(\theta)$ 与分布 $F(x\,|\,\theta)$ 是共轭的.

证明

(a) 如果 $F(\cdot\,|\,\theta)\sim\mathcal{N}(\theta,\,a_0^{-1})$, 并且 $\Pi(\theta)\sim\mathcal{N}(m_0,\,b_0^{-1})$, 则

$$\Pi(\theta\,|\,x_1,\cdots,x_n)\sim\mathcal{N}\left(\frac{b_0m_0+a_0(x_1+\cdots+x_n)}{b_0+na_0},\ \frac{1}{b_0+na_0}\right).$$

(b) 如果 $F(\cdot\,|\,\theta)\sim\mathcal{N}(0,\,\theta^{-1})$, 并且 $\Pi(\theta)\sim\Gamma(k;\,\lambda)$, 即服从密度函数如下的伽玛分布:

$$\gamma_{(k;\,\lambda)}(x)=\frac{\lambda^kx^{k-1}e^{-\lambda x}}{\Gamma(k)}I_{[0,\infty)}(x),$$

其中 $k>0,\ \lambda>0$, 则

$$\Pi(\theta\,|\,x_1,\cdots,x_n)\sim\Gamma\left(k+\frac{1}{2}n;\ \lambda+\frac{1}{2}(x_1^2+\cdots+x_n^2)\right).$$

(c) 如果 $F(\cdot\,|\,\theta)\sim\exp(\theta)$, 即服从参数为 θ 的指数分布, 而 $\Pi(\theta)\sim\Gamma(k;\,\lambda)$, 则

$$\Pi(\theta\,|\,x_1,\cdots,x_n)\sim\Gamma(k+n;\ \lambda+(x_1+\cdots+x_n)).$$

(d) 如果 $F(\cdot\,|\,\theta)\sim\text{Poisson}(\theta)$, 而 $\Pi(\theta)\sim\Gamma(k;\,\lambda)$, 则

$$\Pi(\theta\,|\,x_1,\cdots,x_n)\sim\Gamma(k+(x_1+\cdots+x_n);\ \lambda+n).$$

(e) 如果 $F(\cdot\,|\,\theta)\sim\text{Bernoulli}(\theta)$, 而 $\Pi(\theta)\sim\mathrm{B}(k;\,L)$, 即服从密度函数如下的贝塔分布:

$$\beta_{(k;\,L)}(x)=\frac{x^{k-1}(1-x)^{L-1}}{\mathrm{B}(k;\,L)}I_{(0,1)}(x),$$

则

$$\Pi(\theta\,|\,x_1,\cdots,x_n)\sim\mathrm{B}\left(k+(x_1+\cdots+x_n);\ L+n-(x_1+\cdots+x_n)\right).$$

87. 设随机变量 X 服从下列分布之一: 二项分布, 泊松分布, 几何分布, 负二项分布, 帕雷托分布. 试求事件 $\{X$ 为偶数$\}$ 的概率. (例如, X 服从参数为 p 的几何分布 (参阅《概率》第一卷第 162 页中的表 2), 则 $\mathbf{P}\{X$ 为偶数$\}=\frac{1-p}{2-p}$.)

88. 设随机变量 $\xi_1,\cdots,\xi_n$ 相互独立, 分别服从参数为 $\lambda_1,\cdots,\lambda_n$ 的指数分布.

(a) 证明, $\mathbf{P}\{\xi_1<\xi_2\}=\lambda_1/(\lambda_1+\lambda_2)$.

(b) 证明, 随机变量 $\min\limits_{1\leqslant k\leqslant n}\xi_k$ 服从参数为 $\sum\limits_{k=1}^{n}\lambda_k$ 的指数分布, 并由 (a) 推出:

$$\mathbf{P}\{\xi_j=\min_{1\leqslant k\leqslant n}\xi_k\}=\frac{\lambda_j}{\sum\limits_{k=1}^{n}\lambda_k}.$$

(c) 假设 $\lambda_i\neq\lambda_j,\ i\neq j$, 试求随机变量 $\xi_1+\cdots+\xi_n$ 的分布密度 ($n=2$ 的情形见第 26 题).

(d) 证明, $\mathbf{E}\min(\xi_1,\xi_2)=1/(\lambda_1+\lambda_2)$, 并求 $\mathbf{E}\max(\xi_1,\xi_2)$.

(e) 试求随机变量 $\xi_1-\xi_2$ 的分布密度.

(f) 证明, 随机变量 $\min(\xi_1,\xi_2)$ 与 $\xi_1-\xi_2$ 相互独立.

89. 设随机变量 X 服从密度函数如下的贝塔分布:

$$f(x;\ \alpha,\beta)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1},\quad 0<x<1,$$

其中 $\alpha>0,\beta>0$. 证明,

$$\mathbf{E}X=\frac{\alpha}{\alpha+\beta},\quad \mathbf{D}X=\frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}.$$

§9. 建立具有给定有限维分布的过程

1. 设 $\Omega=[0,1]$, $\mathscr{F}$ 为区间 $[0,1]$ 中的博雷尔集合类, $\mathbf{P}$ 为区间 $[0,1]$ 上的勒贝格测度. 证明, 概率空间 $(\Omega,\mathscr{F},\mathbf{P})$ 在下述意义上是普适的: 对于任何给定的分布函数 $F(x)$, 都可以在 $(\Omega,\mathscr{F},\mathbf{P})$ 上构造出随机变量 $\xi=\xi(\omega),\ \omega\in\Omega$, 使得它的分布函数 $F_\xi(x)=\mathbf{P}\{\xi\leqslant x\}$ 就是 $F(x)$.

提示: 应当令 $\xi(\omega)=F^{-1}(\omega)$, 其中 $F^{-1}(\omega)=\sup\{x:\ F(x)<\omega\},\ 0<\omega<1$; 而 $\xi(0)$ 与 $\xi(1)$ 可以任取.

2. 验证定理 1 和 2 的推论中的分布族的相容性.

3. 试从定理 1 推出定理 2 的推论 4 中的命题.

提示: 请注意, 由 (16) 式所定义的测度形成相容的分布族.

4. 设 F_n 是 $T_n(n\geqslant 1)$ 的分布函数 (见第 4 小节). 证明, $F_{n+1}(t)=\int\limits_0^t F_n(t-s)dF(s),\ n\geqslant 1$, 其中 $F_1=F$.

5. 证明, $\mathbf{P}\{N_t = n\} = F_n(t) - F_{n+1}(t)$ (参阅 (17)).

6. 证明, 第 4 小节中所引入的更新函数 $m(t)$ 满足*更新方程*:

$$m(t) = F(t) + \int_0^t m(t-x)dF(x). \qquad (*)$$

7. 证明, 方程 $(*)$ 在有限区间中的有界函数类中的唯一解是 (20) 式所定义的函数.

8. 设 T 为任一集合.

(a) 假设对每个 $t \in T$, 都给定了某个概率空间 $(\Omega_t, \mathscr{F}_t, \mathbf{P}_t)$. 令 $\Omega = \prod_{t\in T} \Omega_t$, $\mathscr{F} = \prod_{t\in T} \mathscr{F}_t$. 证明, 在 $(\Omega, \mathscr{F})$ 上存在唯一的概率测度 $\mathbf{P}$, 使得如下的独立性成立:

$$\mathbf{P}\left(\prod_{t\in T} B_t\right) = \prod_{t\in T} \mathbf{P}(B_t),$$

其中集合 $B_t \in \mathscr{F}_t$, $t \in T$, 并且对所有的 $t \in T$, 都有 $B_t = \Omega_t$, 包括其个数有限的情形.

提示: 将测度 $\mathbf{P}$ 定义在适当的代数上面, 并运用约内斯库–图尔其定理中的证明方法.

(b) 假设对每个 $t \in T$ 都给定了一个可测空间 $(E_t, \mathscr{E}_t)$ 和它上面的概率 $\mathbf{P}_t$. 证明如下的 (洛姆尼斯基–乌拉姆) 结果: 存在概率空间 $(\Omega, \mathscr{F}, \mathbf{P})$ 和独立随机元 $(X_t)_{t\in T}$, 使得 X_t 为 $\mathscr{F}/\mathscr{E}_t$ 可测, 并且 $\mathbf{P}\{X_t \in B\} = \mathbf{P}_t(B)$, $B \in \mathscr{E}_t$.

§10. 随机变量序列收敛的各种形式

1. 运用定理 5, 证明, 可以把第 6 节中的定理 3 和定理 4 中的几乎必然收敛换为依概率收敛.

提示: 如果 $\xi_n \xrightarrow{\mathbf{P}} \xi$, $|\xi_n| \leqslant \eta$, 而 $\mathbf{E}|\xi_n - \xi| \not\to 0$, 则可找到 $\varepsilon > 0$ 和自然数子列 $(n_k)_{k\geqslant 1}$, 使得 $\mathbf{E}|\xi_{n_k} - \xi| > \varepsilon$. 但是 $\xi_{n_k} \xrightarrow{\mathbf{P}} \xi$, 故由定理 5 知, 存在这样的自然数子列 $(k_l)_{l\geqslant 1}$, 使得 $\xi_{n_{k_l}} \not\to \xi$ ($\mathbf{P}$-a.s.). 接下来再运用第 6 节中的定理 3, 导出与假设 $\mathbf{E}|\xi_n - \xi| \not\to 0$ 的矛盾.

2. 证明, 空间 L^∞ 是完备的.

提示: 在空间 L^∞ 中取基本列, 意即对于 $n \leqslant m$ 和 $a_n \to 0$ $(n\to\infty)$, 满足条件 $\|\xi_m - \xi_n\|_{L^\infty} \leqslant a_n$ 的随机变量序列 $(\xi_k)_{k\geqslant 1}$, 令

$$\xi(\omega) = \begin{cases} \overline{\lim_n}\, \xi_n(\omega), & \text{如果 } \overline{\lim_n}\, \xi_n(\omega) < \infty, \\ 0, & \text{如果 } \overline{\lim_n}\, \xi_n(\omega) = \infty. \end{cases}$$

证明, $\xi(\omega)$ 是随机变量, 且当 $n \to \infty$ 时, 有 $\|\xi - \xi_n\|_{L^\infty} \leqslant a_n \to 0$.

3. 证明, 如果 $\xi_n \xrightarrow{\mathbf{P}} \xi$, 同时 $\xi_n \xrightarrow{\mathbf{P}} \eta$, 则 ξ 与 η 相互等价(意即 $\mathbf{P}\{\xi \neq \eta\} = 0$).

4. 设 $\xi_n \xrightarrow{\mathbf{P}} \xi$, $\eta_n \xrightarrow{\mathbf{P}} \eta$, 并且 ξ 与 η 相互等价. 证明, 对任何 $\varepsilon > 0$, 都有

$$\mathbf{P}\{|\xi_n - \eta_n| \geqslant \varepsilon\} \to 0, \quad n \to \infty.$$

5. 设 $\xi_n \xrightarrow{\mathbf{P}} \xi$, $\eta_n \xrightarrow{\mathbf{P}} \eta$. 证明, 如果 $\varphi = \varphi(x, y)$ 是连续函数, 则有 $\varphi(\xi_n, \eta_n) \xrightarrow{\mathbf{P}} \varphi(\xi, \eta)$ (*斯鲁斯基引理*).

提示: 对任给的 $\varepsilon > 0$, 选取 $c > 0$, 使得

$$\mathbf{P}\{|\xi_n| > c\} < \varepsilon, \quad \mathbf{P}\{|\eta_n| > c\} < \varepsilon, \quad n \geqslant 1,$$
$$\mathbf{P}\{|\xi| > c\} < \varepsilon, \quad \mathbf{P}\{|\eta| > c\} < \varepsilon.$$

由于 $\varphi = \varphi(x, y)$ 连续, 所以它在紧集 $[-c, c]\times[-c, c]$ 上面一致连续. 因此, 存在 $\delta > 0$, 使得只要 $\mathbf{x}, \mathbf{y} \in [-c, c]\times[-c, c]$, 并且 $\rho(\mathbf{x}, \mathbf{y}) < \delta$, 就有 $|\varphi(\mathbf{x}) - \varphi(\mathbf{y})| \leqslant \varepsilon$. (此处, 对于 $\mathbf{x} = (x^{(1)}, x^{(2)})$, $\mathbf{y} = (y^{(1)}, y^{(2)})$, 有 $\rho(\mathbf{x}, \mathbf{y}) = \max(|x^{(1)} - y^{(1)}|, |x^{(2)} - y^{(2)}|)$.) 进而, 再利用

$$\begin{aligned}\mathbf{P}\{|\varphi(\xi_n, \eta_n) - \varphi(\xi, \eta)| \geqslant \varepsilon\} &\leqslant \mathbf{P}\{|\xi_n| > c\} + \mathbf{P}\{|\eta_n| > c\} \\ &+ \mathbf{P}\{|\xi| > c\} + \mathbf{P}\{|\eta| > c\} + \mathbf{P}\{|\xi_n - \xi| > \delta\} + \mathbf{P}\{|\eta_n - \eta| > \delta\},\end{aligned}$$

即可证得, 对于充分大的 n, 上式右端小于 6ε.

6. 设 $(\xi_n - \xi)^2 \xrightarrow{\mathbf{P}} 0$. 证明 $\xi_n^2 \xrightarrow{\mathbf{P}} \xi^2$.

7. 证明, 如果 $\xi_n \xrightarrow{\mathrm{d}} C$, 其中 C 为常数, 则成立依概率收敛性, 亦即

$$\xi_n \xrightarrow{\mathrm{d}} C \quad\Rightarrow\quad \xi_n \xrightarrow{\mathbf{P}} C.$$

提示: 任取 $\varepsilon > 0$, 考察函数 $f_\varepsilon(x) = \left(1 - \frac{|x-C|}{\varepsilon}\right)^+$. 我们有

$$\mathbf{P}\{|\xi_n - C| \geqslant \varepsilon\} \geqslant \mathbf{E} f_\varepsilon(\xi_n) \to \mathbf{E} f_\varepsilon(C) = 1.$$

8. 设 $(\xi_n)_{n\geqslant 1}$ 为随机变量序列, 对某 $p > 0$, 有 $\sum_{n=1}^\infty \mathbf{E}|\xi_n|^p < \infty$. 证明, $\xi_n \to 0$ ($\mathbf{P}$-a.s.)

提示: 利用切比雪夫不等式和博雷尔–坎泰利引理.

9. 设 $(\xi_n)_{n\geqslant 1}$ 为独立同分布随机变量序列. 证明如下各蕴涵关系:

$$\mathbf{E}|\xi_1| < \infty \quad\Leftrightarrow\quad \sum_{n=1}^\infty \mathbf{P}\{|\xi_1| > \varepsilon n\} < \infty, \quad \varepsilon > 0$$
$$\Leftrightarrow\quad \sum_{n=1}^\infty \mathbf{P}\left\{\left|\frac{\xi_n}{n}\right| > \varepsilon\right\} < \infty, \quad \varepsilon > 0 \quad\Rightarrow\quad \frac{\xi_n}{n} \to 0 \quad (\mathbf{P}\text{-a.s.}).$$

提示: 利用下面的易于验证的不等式

$$\varepsilon\sum_{n=1}^{\infty}\mathbf{P}\{|\xi_1|>\varepsilon n\}\leqslant\mathbf{E}|\xi_1|\leqslant\varepsilon+\varepsilon\sum_{n=1}^{\infty}\mathbf{P}\{|\xi_1|>\varepsilon n\}.$$

10. 设 $(\xi_n)_{n\geqslant1}$ 为随机变量序列, ξ 为某个随机变量.

(a) 假设对每个 $\varepsilon>0$, 都有 $\mathbf{P}\{|\xi_n-\xi|\geqslant\varepsilon\ \text{i.o.}\}=0$. 证明, $\xi_n\to\xi$ ($\mathbf{P}$-a.s.).

(b) 假设存在子列 (n_k), 使得当 $k\to\infty$ 时, 有 $\xi_{n_k}\to\xi$ ($\mathbf{P}$-a.s.) 和 $\max\limits_{n_{k-1}<l\leqslant n_k}|\xi_l-\xi_{n_{k-1}}|\to0$ ($\mathbf{P}$-a.s.). 证明, $\xi_n\to\xi$ ($\mathbf{P}$-a.s.).

(c) 假设 $\xi_n\to\xi$ ($\mathbf{P}$-a.s.). 证明, 此时 (a) 的逆命题成立, 即对任何 $\varepsilon>0$, 都有 $\mathbf{P}\{|\xi_n-\xi|\geqslant\varepsilon\ \text{i.o.}\}=0$.

11. 在所有随机变量的集合中定义 "d 距离", 令

$$d(\xi,\eta)=\mathbf{E}\frac{|\xi-\eta|}{1+|\xi-\eta|},$$

并将几乎必然相等的随机变量视为相同. 证明, "d 距离" 的确是一种距离, 并证明, 依概率收敛等价于按该种距离收敛.

提示: 验证三角形不等式, 并对任何 $\varepsilon>0$, 证明下述不等式:

$$\frac{\varepsilon}{1+\varepsilon}\mathbf{P}\{|\xi_n-\xi|>\varepsilon\}\leqslant\mathbf{E}\frac{|\xi_n-\xi|}{1+|\xi_n-\xi|}\leqslant\varepsilon\mathbf{P}\{|\xi_n-\xi|\leqslant\varepsilon\}+\mathbf{P}\{|\xi_n-\xi|>\varepsilon\}.$$

12. 证明, 在所有随机变量的集合中不存在这样的距离, 使得按该种距离收敛等价于几乎必然收敛.

提示: 假设存在某种距离 ρ, 使得按其收敛蕴涵几乎必然收敛, 我们来考察这样的随机变量序列 $(\xi_n)_{n\geqslant1}$: 对其有 $\xi_n\xrightarrow{\mathbf{P}}0$ 和 $\xi_n\not\to0$ ($\mathbf{P}$-a.s.). 此时显然可以找到某个 $\varepsilon>0$ 和某个子列 $(n_k)_{k\geqslant1}$, 使得 $\rho(\xi_{n_k},0)>\varepsilon$, 但是却有 $\xi_{n_k}\xrightarrow{\mathbf{P}}0$. 再由定理 5, 即可导出与蕴涵几乎必然收敛的距离 ρ 的存在性的矛盾.

13. 设 $X_1\leqslant X_2\leqslant\cdots$, 且 $X_n\xrightarrow{\mathbf{P}}X$. 证明, 此时我们有 $X_n\to X$ ($\mathbf{P}$-a.s.).

14. 设 $(X_n)_{n\geqslant1}$ 为随机变量序列. 证明,

(a) $X_n\to0$ ($\mathbf{P}$-a.s.) $\Rightarrow$ $S_n/n\to0$ ($\mathbf{P}$-a.s.), 其中 $S_n=X_1+\cdots+X_n$.

(b) 如果 $p\geqslant1$, 则 $X_n\xrightarrow{L^p}0\ \Rightarrow\ S_n/n\xrightarrow{L^p}0$, 但是一般来说, 有

$$X_n\xrightarrow{L^p}0\ \not\Leftarrow\ \frac{S_n}{n}\xrightarrow{L^p}0.$$

(c) $X_n\xrightarrow{\mathbf{P}}0$, 一般来说, 不能蕴涵 $S_n/n\xrightarrow{\mathbf{P}}0$ (试比较第 34 题中的断言).

15. 设 $(\Omega,\mathscr{F},\mathbf{P})$ 为概率空间, 且 $X_n\xrightarrow{\mathbf{P}}X$. 证明, 如果测度 $\mathbf{P}$ 是原子的, 则亦概率为 1 地有 $X_n\to X$. (参阅第二章第 3 节第 35 题关于原子概率测度的定义.)

16. 根据博雷尔–坎泰利引理中的断言 a)(亦称为博雷尔–坎泰利第一引理), 如果对任何 $\varepsilon>0$, 都有 $\sum\limits_{n=1}^{\infty}\mathbf{P}\{|\xi_n|>\varepsilon\}<\infty$, 则有 $\xi_n\to0$ ($\mathbf{P}$-a.s.). 试举例说明, 存在随机变量 $\xi_n\to0$ ($\mathbf{P}$-a.s.), 然而却有 $\sum\limits_{n=1}^{\infty}\mathbf{P}\{|\xi_n|>\varepsilon\}=\infty$, $\varepsilon>0$.

17. (关于博雷尔–坎泰利引理中的断言 b) (亦称为博雷尔–坎泰利第二引理)) 设 $\Omega=(0,1)$, $\mathscr{F}=\mathscr{B}((0,1))$, $\mathbf{P}$ 为勒贝格测度. 我们来考察事件 $A_n=(0,1/n)$. 证明, $\sum\limits_{n=1}^{\infty}\mathbf{P}(A_n)=\infty$, 但是对于区间 $(0,1)$ 中的每个 ω, 都只可能属于有限个集合 $A_1,\cdots,A_{[1/\omega]}$, 亦即有 $\mathbf{P}\{A_n\ \text{i.o.}\}=0$.

18. 证明, 为了使得博雷尔–坎泰利引理中的断言 b) 即第二引理中的断言成立, 可以不必要求事件序列的独立性, 而只需要求事件 $A_1,A_2,\cdots$ 两两独立 (亦即 $\mathbf{P}(A_i\cap A_j)-\mathbf{P}(A_i)\mathbf{P}(A_j)=0$, $i\neq j$) 或两两负相关 (亦即 $\mathbf{P}(A_i\cap A_j)-\mathbf{P}(A_i)\mathbf{P}(A_j)\leqslant0$, $i\neq j$) 即可.

19. (关于博雷尔–坎泰利第二引理) 证明博雷尔–坎泰利第二引理的如下形式: 对于任何 (不一定相互独立的) 事件序列 $A_1,A_2,\cdots$, 有

(a) 如果

$$\sum_{n=1}^{\infty}\mathbf{P}(A_n)=\infty\quad\text{且}\quad\liminf_n\frac{\sum\limits_{i,k=1}^{n}\mathbf{P}(A_iA_k)}{\left[\sum\limits_{k=1}^{n}\mathbf{P}(A_k)\right]^2}=1,$$

则 $\mathbf{P}\{A_n\ \text{i.o.}\}=1$(爱尔迪希 (Erdös), 雷尼 (Rényi)).

(b) 如果

$$\sum_{n=1}^{\infty}\mathbf{P}(A_n)=\infty\quad\text{且}\quad\liminf_n\frac{\sum\limits_{i,k=1}^{n}\mathbf{P}(A_iA_k)}{\left[\sum\limits_{k=1}^{n}\mathbf{P}(A_k)\right]^2}=L,$$

则 $L\geqslant1$, 并且 $\mathbf{P}\{A_n\ \text{i.o.}\}=1/L$(科琴 (Kochen), 斯通 (Stone)[65]; 斯皮策 (Spitzer) [113]).

(c) 如果

$$\sum_{n=1}^{\infty}\mathbf{P}(A_n)=\infty\quad\text{且}\quad\liminf_n\frac{\sum\limits_{1\leqslant i<k\leqslant n}[\mathbf{P}(A_iA_k)-\mathbf{P}(A_i)\mathbf{P}(A_k)]}{\left[\sum\limits_{k=1}^{n}\mathbf{P}(A_k)\right]^2}\leqslant0,$$

则 $\mathbf{P}\{A_n\ \text{i.o.}\}=1$(奥尔特加 (Ortega), 弗舍波尔 (Wschebor) [86]).

(d) 如果 $\sum_{n=1}^{\infty} \mathbf{P}(A_n) = \infty$, 并且

$$\alpha_H = \liminf_n \frac{\sum_{1\leqslant i<k\leqslant n} [\mathbf{P}(A_iA_k) - H\mathbf{P}(A_i)\mathbf{P}(A_k)]}{\left[\sum_{k=1}^{n} \mathbf{P}(A_k)\right]^2},$$

其中 H 为任一常数, 则都有 $\mathbf{P}\{A_n \text{ i.o.}\} \geqslant \frac{1}{H+2\alpha_H}$, 其中, $H+2\alpha_H \geqslant 1$ (彼得罗夫 (Petrov) [87]).

20. 设 $A_1, A_2, \cdots$ 为相互独立的随机事件序列, 有 $\sum_{n=1}^{\infty} \mathbf{P}(A_n) = \infty$. 证明, 对 $S_n = \sum_{k=1}^{n} I(A_k)$, 有如下的加强的博雷尔–坎泰利第二引理成立:

$$\lim_n \frac{S_n}{\mathbf{E}S_n} = 1 \quad (\mathbf{P}\text{-a.s.}).$$

21. 设 $(X_n)_{n\geqslant 1}$ 与 $(Y_n)_{n\geqslant 1}$ 为两个随机变量序列, 它们的所有有限维分布都对应相等 ($F_{X_1,\cdots,X_n} = F_{Y_1,\cdots,Y_n}$, $n \geqslant 1$). 假设 $X_n \xrightarrow{\mathbf{P}} X$. 证明, 存在某个与 X 同分布的随机变量 Y (即有 Law (X) =Law (Y) 或写为 $X \overset{\mathrm{d}}{=} Y$, $X \overset{\mathrm{law}}{=} Y$), 使得 $Y_n \xrightarrow{\mathbf{P}} Y$.

22. 设 $(X_n)_{n\geqslant 1}$ 为独立随机变量序列, 对于某个随机变量 X, 有 $X_n \xrightarrow{\mathbf{P}} X$. 证明, X 为退化随机变量.

23. 证明, 对于任何随机变量序列 $\xi_1, \xi_2, \cdots$, 都能找到常数序列 $a_1, a_2, \cdots$, 使得 $\xi_n/a_n \to 0$ ($\mathbf{P}$-a.s.).

24. 设 $\xi_1, \xi_2, \cdots$ 为随机变量序列, 而 $S_n = \xi_1 + \cdots + \xi_n$, $n \geqslant 1$. 证明, 集合 $(S_n \to)$(由所有使得 $\sum_{k\geqslant 1} \xi_k(\omega)$ 收敛的 $\omega \in \Omega$ 构成的集合) 可以表示为

$$(S_n \to) = \bigcap_{N\geqslant 1} \bigcup_{m\geqslant 1} \bigcap_{k\geqslant m} \left\{ \sup_{l\geqslant k} |S_l - S_k| \leqslant N^{-1} \right\}.$$

相应的, 由使得 $\sum_{k\geqslant 1} \xi_k(\omega)$ 不收敛的 $\omega \in \Omega$ 所构成的集合 $(S_n \not\to)$ 可以表示为

$$(S_n \not\to) = \bigcup_{N\geqslant 1} \bigcap_{m\geqslant 1} \bigcup_{k\geqslant m} \left\{ \sup_{l\geqslant k} |S_l - S_k| > N^{-1} \right\}.$$

25. 设 Ω 为至多可数集. 证明, 此时若有 $\xi_n \xrightarrow{\mathbf{P}} \xi$, 则有 $\xi_n \to \xi$ ($\mathbf{P}$-a.s.).

26. 试举例说明, 存在随机变量序列 $(\xi_n)_{n\geqslant 1}$, 对其概率为 1 地有 $\limsup \xi_n = \infty$, $\liminf \xi_n = -\infty$, 但是却存在随机变量 η, 使得 $\xi_n \xrightarrow{\mathbf{P}} \eta$.

27. 证明 0-1 律的下述形式 (试比较第四章第 1 节中的 0-1 律): 如果 $A_1, A_2, \cdots$ 为两两独立的随机事件序列, 则

$$\mathbf{P}\{A_n \text{ i.o.}\} = \begin{cases} 0, & \text{如果 } \sum \mathbf{P}(A_n) < \infty, \\ 1, & \text{如果 } \sum \mathbf{P}(A_n) = \infty. \end{cases}$$

28. 设 $A_1, A_2, \cdots$ 为任一随机事件序列, 有 $\lim_n \mathbf{P}(A_n) = 0$ 和 $\sum_n \mathbf{P}(A_n \cap \overline{A}_{n+1}) < \infty$. 证明, 此时有 $\mathbf{P}\{A_n \text{ i.o.}\} = 0$.

29. 证明, 如果 $\sum_n \mathbf{P}\{|\xi_n| > n\} < \infty$, 则 $\limsup_n (|\xi_n|/n) \leqslant 1$ ($\mathbf{P}$-a.s.).

30. 设 $\xi_n \downarrow \xi$ ($\mathbf{P}$-a.s.), $\mathbf{E}|\xi_n| < \infty$, $n \geqslant 1$ 且 $\inf_n \mathbf{E}\xi_n > -\infty$. 证明 $\xi_n \xrightarrow{L^1} \xi$, 亦即 $\mathbf{E}|\xi_n - \xi| \to 0$.

31. 利用博雷尔–坎泰利第二引理, 证明, $\mathbf{P}\{A_n \text{ i.o.}\} = 1$, 当且仅当, 对任何使得 $\mathbf{P}(A) > 0$ 的集合 A, 都有 $\sum_n \mathbf{P}(A \cap A_n) = \infty$.

32. 设随机事件 $A_1, A_2, \cdots$ 相互独立, 且对一切 $n \geqslant 1$, 都有 $\mathbf{P}(A_n) < 1$. 证明 $\mathbf{P}\{A_n \text{ i.o.}\} = 1$, 当且仅当, $\mathbf{P}(\cup A_n) = 1$.

33. 设 $X_1, X_2, \cdots$ 为独立随机变量序列, 有 $\mathbf{P}\{X_n = 0\} = 1/n$, $\mathbf{P}\{X_n = 1\} = 1-1/n$. 令 $E_n = \{X_n = 0\}$. 试由性质 $\sum_{n=1}^{\infty} \mathbf{P}(E_n) = \infty$ 与 $\sum_{n=1}^{\infty} \mathbf{P}(\overline{E}_n) = \infty$, 推出极限 $\lim_n X_n$ ($\mathbf{P}$-a.s.) 不存在.

34. 设 $X_1, X_2, \cdots$ 为随机变量序列. 证明 $X_n \xrightarrow{\mathbf{P}} 0$, 当且仅当, 对某个 $r > 0$, 有

$$\mathbf{E}\frac{|X_n|^r}{1+|X_n|^r} \to 0.$$

特别地, 如果 $S_n = X_1 + \cdots + X_n$, 则有

$$\frac{S_n - \mathbf{E}S_n}{n} \xrightarrow{\mathbf{P}} 0 \quad \Leftrightarrow \quad \mathbf{E}\frac{(S_n - \mathbf{E}S_n)^2}{n^2 + (S_n - \mathbf{E}S_n)^2} \to 0.$$

再证明如下的蕴涵关系:

$$\max_{1\leqslant k\leqslant n} |X_k| \xrightarrow{\mathbf{P}} 0 \quad \Rightarrow \quad \frac{S_n}{n} \xrightarrow{\mathbf{P}} 0, \ n \to \infty.$$

35. 设 $X_1, X_2, \cdots$ 为独立同分布随机变量序列, 有 $\mathbf{P}\{X_k = \pm 1\} = 1/2$. 令 $U_n = \sum_{k=1}^{n} \frac{X_k}{2^k}$, $n \geqslant 1$. 证明, $U_n \to U$ ($\mathbf{P}$-a.s.), 其中随机变量 U 服从区间 $[-1, +1]$ 上的均匀分布.

36. (叶戈罗夫定理) 设 $(\Omega, \mathscr{F}, \mu)$ 为赋有有限测度 μ 的可测空间, 而 $f_1, f_2, \cdots$ 为定义在其上的 μ-几乎处处收敛于博雷尔函数 f 的博雷尔函数序列 ($f_n \xrightarrow{\mu\text{-a.s.}} f$). 叶

戈罗夫 (Egorov) 定理断言: 对任何给定的 $\varepsilon > 0$, 都存在 $A_\varepsilon \in \mathscr{F}$, 有 $\mu(A_\varepsilon) < \varepsilon$, 使得在其余集 $\overline{A}_\varepsilon = \Omega \backslash A_\varepsilon$ 上面, 成立一致收敛性 $f_n(\omega) \to f(\omega)$, $\omega \in \overline{A}_\varepsilon$. 试证明该断言.

37. (鲁金定理) 设 $(\Omega, \mathscr{F}, \mathbf{P}) = ([a,b], \mathscr{F}, \lambda)$, 其中 λ 是区间 $[a,b]$ 上的勒贝格测度, $\mathscr{F}$ 为勒贝格集合类. 设 $f = f(x)$ 是勒贝格可测的有限函数. 试证明鲁金定理: 对任何 $\varepsilon > 0$, 都能找到连续函数 $f_\varepsilon = f_\varepsilon(x)$, 使得

$$\lambda\{x \in [a,b]:\ f(x) \neq f_\varepsilon(x)\} < \varepsilon.$$

38. 叶戈罗夫定理自然地导致如下概念的引入:

称函数序列 $f_1, f_2, \cdots$ 几乎一致地收敛到函数 f, 如果对任何 $\varepsilon > 0$, 都存在 $A_\varepsilon \in \mathscr{F}$, $\mu(A_\varepsilon) < \varepsilon$, 使得 $f_n(\omega) \to f(\omega)$ 对一切 $\omega \in \overline{A}_\varepsilon$ 一致成立 (记为 $f_n \rightrightarrows f$).

证明, 几乎一致收敛 $f_n \rightrightarrows f$ 蕴涵依测度收敛 $(f_n \xrightarrow{\mu} f)$ 和几乎必然收敛 $(f_n \xrightarrow{\mu\text{-a.s.}} f)$.

39. 设 $X_1, X_2, \cdots$ 为随机变量序列, 而 $\{X_n \not\to\}$ 为当 $n \to \infty$ 时, 那些使得 $X_n(\omega)$ 不收敛的 $\omega \in \Omega$ 所构成的集合. 证明

$$\{X_n \not\to\} = \bigcup_{p<q} \{\liminf X_n \leqslant p < q \leqslant \limsup X_n\},$$

其中的求并运算系对一切满足 $p < q$ 的有理数 p 和 q 所构成的数对 (p,q) 求并.

40. 设随机变量序列 $X_1, X_2, \cdots$ 定义在完备的概率空间之上, 并且概率为 1 地收敛到随机变量 X. 证明, 分别由随机元 $(X_1, X_2, \cdots)$ 与 $(X_1, X_2, \cdots, X)$ 所生成的 σ-代数 (参阅第 5 节) $\sigma(X_1, X_2, \cdots)$ 与 $\sigma(X_1, X_2, \cdots, X)$ 相互重合.

41. 设 $X_1, X_2, \cdots$ 为独立同分布随机变量序列, 它们的共同分布 $F = F(x)$ 满足条件

$$\lim_{x\to\infty} x^2[1 - F(x)] = 0.$$

证明

$$\frac{1}{\sqrt{n}} \max_{1\leqslant i\leqslant n} X_i \xrightarrow{\mathbf{P}} 0, \quad n \to \infty.$$

42. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布随机变量序列, 有 $\mathbf{E}\xi_1 = \mu$, $\mathbf{D}\xi_1 = \sigma^2 < \infty$ 和 $\mathbf{P}\{\xi_1 = 0\} = 0$. 证明,

$$\frac{\sum\limits_{k=1}^n \xi_k}{\sum\limits_{i=1}^n \xi_k^2} \xrightarrow{\mathbf{P}} \frac{\mu}{\mu^2 + \sigma^2}, \quad n \to \infty.$$

43. 假设 $\xi_n \xrightarrow{\mathbf{P}} \xi$, $\eta_n \xrightarrow{\mathbf{P}} \eta$, 并且有 $\mathbf{P}\{\xi_n \leqslant \eta_n\} = 1$, $n \geqslant 1$. 证明, $\mathbf{P}\{\xi \leqslant \eta\} = 1$.

44. 设 $\xi_1, \xi_2, \cdots$ 为非负随机变量序列, $\mathscr{F}_1, \mathscr{F}_2, \cdots$ 为 σ-代数序列, 有 $\mathbf{E}(\xi_n \,|\, \mathscr{F}_n) \xrightarrow{\mathbf{P}} 0$. 证明, $\xi_n \xrightarrow{\mathbf{P}} 0$.

45. 假设 $\xi_n \xrightarrow{\text{d}} \xi$, 且 $c_n\xi_n \xrightarrow{\text{d}} \xi$ ($\xrightarrow{\text{d}}$ 表示依分布收敛), 其中 ξ 为非退化随机变量, $c_n > 0$. 证明, $c_n \to 1$.

46. 设 $X_1, X_2, \cdots$ 为任一随机变量序列. 证明

(a) 若 $\overline{\lim} \frac{(\mathbf{E}X_n)^2}{\mathbf{E}X_n^2} = 1$, 则 $\lim \frac{X_n}{\mathbf{E}X_n} = 0$ ($\mathbf{P}$-a.s.)(试比较第 20 题中的断言).

(b) 若 $\overline{\lim} \frac{(\mathbf{E}X_n)^2}{\mathbf{E}X_n^2} > 0$, $\mathbf{E}X_n \neq 0$, $0 < \mathbf{E}X_n^2 < \infty$, 则

$$\mathbf{P}\left(\underline{\lim} \frac{X_n}{\mathbf{E}X_n} \leqslant 1 \leqslant \overline{\lim} \frac{X_n}{\mathbf{E}X_n}\right) > 0.$$

47. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布随机变量序列, 有 $\mathbf{E}|\xi_1| < \infty$. 设 α 为某个正的常数, 令 $A_n = \{|\xi_n| > \alpha n\}$, $n \geqslant 1$. 证明, $\mathbf{P}(\overline{\lim} A_n) = 0$.

48. 证明, 在连续函数空间 C 中不存在这样的距离 ρ, 使得收敛性 $\rho(f_n, f) \to 0$ 等价于 f_n 逐点收敛到 f(试比较第 12 题).

49. 试举例说明, 对任何 $c > 0$, 都存在随机变量序列 $\xi, \xi_1, \xi_2, \cdots$, 使得逐点地有 $\xi_n(\omega) \to \xi(\omega)$, 并且 $\mathbf{E}\xi_n = -c$, $n \geqslant 1$, 但是却有 $\mathbf{E}\xi = c$.

50. 试 (在第一章第 4 节第 23 题中的三种定义中的每一种意义之下) 求出随机变量 $\xi_n = I_{\{\eta \leqslant (-1)^n - 1/n\}}$ 的中位数 $\mu_n = \mu(\xi_n)$, 其中 η 是标准高斯随机变量.

51. 设随机变量 $\xi_n (n \geqslant 1)$ 的中位数 $\mu_n = \mu(\xi_n)$ 唯一确定 (参阅第一章第 4 节第 23 题), ξ_n 几乎必然收敛到随机变量 ξ. 试举例说明, 一般来说, 极限 $\lim\limits_n \mu_n$ 未必存在.

52. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布的非负随机变量序列, 有 $\mathbf{E}\xi_1 = 1$. 令 $T_n = \prod\limits_{k=1}^n \xi_k$, $n \geqslant 1$. 证明, 当 $n \to \infty$ 时, 有 $T_n \to 0$ ($\mathbf{P}$-a.s.).

53. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布随机变量序列. 令 $S_n = \xi_1 + \cdots + \xi_n$, $n \geqslant 1$. 证明,

(a) $\mathbf{E}\xi_1^+ = \infty$, $\mathbf{E}\xi_1^- < \infty \quad \Rightarrow \quad \frac{S_n}{n} \to +\infty$ ($\mathbf{P}$-a.s.).

(b) $\mathbf{E}|\xi_1| < \infty$, $\mathbf{E}\xi_1 \neq 0 \quad \Rightarrow \quad \dfrac{\max\,(|\xi_1|, \cdots, |\xi_n|)}{|S_n|} \to 0$ ($\mathbf{P}$-a.s.).

(c) $\mathbf{E}\xi_1^2 < \infty \quad \Rightarrow \quad \dfrac{\max\,(|\xi_1|, \cdots, |\xi_n|)}{\sqrt{n}} \to 0$ ($\mathbf{P}$-a.s.).

54. 设 $\xi_1, \xi_2, \cdots$ 为随机变量序列, 而 ξ 为某个随机变量. 证明, 对一切 $p \geqslant 1$, 都有如下两个条件相互等价:

a) $\xi_n \xrightarrow{L^p} \xi$(亦即 $\mathbf{E}|\xi_n - \xi|^p \to 0$).

b) $\xi_n \xrightarrow{\mathbf{P}} \xi$, 并且随机变量族 $\{|\xi_n|^p,\ n \geqslant 1\}$ 一致可积.

55. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布随机变量序列, 有 $\mathbf{P}\{\xi_1 > x\} = e^{-x}$, $x \geqslant 0$ (即服从

标准指数分布). 证明

$$\mathbf{P}\{\xi_n > \alpha \ln n \ \text{ i.o.}\} = \begin{cases} 1, & 如果\ \alpha \leqslant 1, \\ 0, & 如果\ \alpha > 1. \end{cases}$$

上述断言还可加强为

$$\mathbf{P}\{\xi_n > \ln n + \alpha \ln\ln n \ \text{ i.o.}\} = \begin{cases} 1, & 如果\ \alpha \leqslant 1, \\ 0, & 如果\ \alpha > 1. \end{cases}$$

并且, 一般地, 对任何 $k \geqslant 1$, 都有

$$\mathbf{P}\{\xi_n > \ln n + \ln\ln n + \cdots + \underbrace{\ln\cdots\ln}_{k\,重} n + +\alpha \underbrace{\ln\cdots\ln}_{(k+1)\,重} n \ \text{ i.o.}\} = \begin{cases} 1, & 如果\ \alpha \leqslant 1, \\ 0, & 如果\ \alpha > 1. \end{cases}$$

56. 为推广定理 3 (关于由几乎必然收敛推出在 L^1 中的收敛性), 试证明如下结果 (谢费引理): 设 ξ 为随机变量, 而 $(\xi_n)_{n\geqslant 1}$ 为随机变量序列, 有 $\mathbf{E}|\xi| < \infty$, $\mathbf{E}|\xi_n| < \infty$ 和 $\xi_n \to \xi$ ($\mathbf{P}$-a.s.), 则 $\mathbf{E}|\xi_n - \xi| \to 0$, 当且仅当, $\mathbf{E}|\xi_n| \to \mathbf{E}|\xi|$, $n \to \infty$. (试比较第 6 节第 19 题中的结论.)

57. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布随机变量序列, 具有密度函数 $f = f(x)$. 现知

$$\lim_{x \downarrow 0} f(x) = \lambda > 0.$$

证明

$$n \min(\xi_1, \cdots, \xi_n) \xrightarrow{\mathbf{P}} \eta,$$

其中 η 是以 λ 为参数的指数分布随机变量.

58. 证明, 只要第 6 节第 33 题中的条件 (i), (ii), (iii) 之一成立, 就有

$$\mathbf{E}|X_n|^p \to \mathbf{E}|X|^p \quad 对一切\ 0 < p < r.$$

59. 设 $(\xi_n)_{n\geqslant 1}$ 为相互独立的正态随机变量, 有 $\xi_n \sim \mathcal{N}(\mu_n, \sigma_n^2)$. 证明, $\sum\limits_{n\geqslant 1} \xi_n^2$ 在 L^1 中收敛, 当且仅当

$$\sum_{n\geqslant 1} (\mu_n^2 + \sigma_n^2) < \infty.$$

再证明, 当该条件成立时, 对一切 $p \geqslant 1$, 级数 $\sum\limits_{n\geqslant 1} \xi_n^2$ 都在 L^p 中收敛.

提示: 为证第二个结论, 最好是对任何 $p \geqslant 1$, 验证

$$\Big\| \sum_{n\geqslant 1} \xi_n^2 \Big\|_p < \infty.$$

60. 设 $X_1, X_2, \cdots$ 为独立随机变量序列, 均服从区间 $[0,1]$ 上的均匀分布. 令 $Y_n = X_1 \cdots X_n$, $n \geqslant 1$, 我们来考察 $\sum\limits_{n=1}^{\infty} z^n Y_n$. 证明该级数的收敛半径 $R = R(\omega)$ 概率为 1 地等于 e.

提示: 利用 $1/R = \overline{\lim\limits_n} |Y_n|^{1/n}$.

61. 设 $(\Omega, \mathscr{F}, \mathbf{P}) = ([0,1), \mathscr{B}([0,1)), \lambda)$, 其中 λ 为勒贝格测度. 设实数 $\omega \in [0,1)$, 而 $\omega = (a_1, a_2, \cdots)$ 是它的连分数分解 ($a_n = a_n(\omega)$ 为整数) [3]. 证明, 当 $n \to \infty$ 时, 有

$$\lambda\,\{\omega:\ a_n(\omega) = k\} \to \frac{1}{\ln 2} \ln \left[\frac{1 + 1/k}{1 + 1/(k+1)} \right].$$

注: 在 "概率的数学理论形成的简史"(《概率》第二卷 p313) 和 V. I. 阿诺尔德的专著 [4] 的 p101 中可以找到关于该问题的提出和解决过程的介绍.

62. 设 $X_1, X_2, \cdots$ 为独立同分布随机变量序列, 有 $\mathbf{P}\{X_1 = 0\} = \mathbf{P}\{X_1 = 2\} = 1/2$. 证明

(a) 级数 $\sum\limits_{n=1}^{\infty} \frac{X_n}{3^n}$ 几乎必然收敛 (到某个随机变量 X).

(b) 随机变量 X 的分布函数是康托尔函数 (参阅第 3 节).

§11. 具有有限二阶矩的随机变量的希尔伯特空间

1. 证明, 如果 $\xi = \text{l.i.m.}\ \xi_n$, 则 $\|\xi_n\| \to \|\xi\|$.

2. 证明, 如果 $\xi = \text{l.i.m.}\ \xi_n$ 且 $\eta = \text{l.i.m.}\ \eta_n$, 则 $(\xi_n, \eta_n) \to (\xi, \eta)$.

3. 证明, 范数 $\|\cdot\|$ 满足 "平行四边形" 性质:

$$\|\xi + \eta\|^2 + \|\xi - \eta\|^2 = 2(\|\xi\|^2 + \|\eta\|^2).$$

4. 设 $\{\xi_1, \cdots, \xi_n\}$ 为正交随机变量族. 证明, 对其成立 "毕达哥拉斯定理":

$$\Big\| \sum_{i=1}^{n} \xi_i \Big\|^2 = \sum_{i=1}^{n} \|\xi_i\|^2.$$

5. 设 $\xi_1, \xi_2, \cdots$ 为正交随机变量序列, 令 $S_n = \xi_1 + \cdots + \xi_n$. 证明, 如果 $\sum\limits_{n=1}^{\infty} \mathbf{E}\xi_n^2 < \infty$, 则存在随机变量 S, 有 $\mathbf{E}S^2 < \infty$, 使得 $\text{l.i.m.}\ S_n = S$, 亦即 $\|S_n - S\|^2 = \mathbf{E}\|S_n - S\|^2 \to 0$, $n \to \infty$.

提示: 根据第 4 题, 利用关系式 $\|S_{n+k} - S_n\|^2 = \sum\limits_{m=n+1}^{n+k} \|\xi_m\|^2$.

6. 证明, 拉德马赫 (Rademacher) 函数 R_n 可按如下方式定义:

$$R_n(x) = \operatorname{sign}(\sin 2^n \pi x), \quad 0 \leqslant x \leqslant 1, \quad n = 1, 2, \cdots.$$

7. 证明, 对于 $\xi \in L^2(\Omega, \mathscr{F}, \mathbf{P})$ 和任何子 σ-代数 $\mathscr{G} \subseteq \mathscr{F}$, 都有

$$\|\xi\| \geqslant \|\mathbf{E}(\xi \mid \mathscr{G})\|,$$

并且等号成立, 当且仅当, $\xi = \mathbf{E}(\xi \mid \mathscr{G})$ ($\mathbf{P}$-a.s.).

8. 证明, 如果 $X, Y \in L^2(\Omega, \mathscr{F}, \mathbf{P})$, $\mathbf{E}(X \mid Y) = Y$, $\mathbf{E}(Y \mid X) = X$, 则有 $X = Y$ ($\mathbf{P}$-a.s.). (试比较第 7 节第 24 题中关于 $X, Y \in L^1(\Omega, \mathscr{F}, \mathbf{P})$ 时的断言, 那里的证明要比 $X, Y \in L^2(\Omega, \mathscr{F}, \mathbf{P})$ 场合下的证明复杂得多.)

9. 给定 $\mathscr{F}$ 的三个子 σ-代数 $(\mathscr{G}_n^{(1)})$, $(\mathscr{G}_n^{(2)})$ 和 $(\mathscr{G}_n^{(3)})$. 设 ξ 为有界随机变量. 假设对每个 n, 都有

$$\mathscr{G}_n^{(1)} \subseteq \mathscr{G}_n^{(2)} \subseteq \mathscr{G}_n^{(3)}, \quad \mathbf{E}(\xi \mid \mathscr{G}_n^{(1)}) \xrightarrow{\mathbf{P}} \eta, \quad \mathbf{E}(\xi \mid \mathscr{G}_n^{(3)}) \xrightarrow{\mathbf{P}} \eta.$$

证明, $\mathbf{E}(\xi \mid \mathscr{G}_n^{(2)}) \xrightarrow{\mathbf{P}} \eta$.

10. 设 $f = f(x)$ 为定义在 $[0, \infty)$ 上的博雷尔函数, 对任何 $\lambda > 0$, 都有

$$\int_0^\infty e^{-\lambda x} f(x) dx = 0.$$

证明 (关于勒贝格测度几乎必然) $f = 0$.

11. 设随机变量 η 服从区间 $[-1, 1]$ 上的均匀分布, 而 $\xi = \eta^2$. 证明,

(a) 根据 η 估计 ξ 和根据 ξ 估计 η (在均方差意义下) 的最优估计分别由下面的关系式给出:

$$\mathbf{E}(\xi \mid \eta) = \eta^2, \quad \mathbf{E}(\eta \mid \xi) = 0.$$

(b) 相应的最优线性估计则分别为:

$$\widehat{\mathbf{E}}(\xi \mid \eta) = 1/3, \quad \widehat{\mathbf{E}}(\eta \mid \xi) = 0.$$

§12. 特征函数

1. 设 ξ 与 η 为相互独立的随机变量, $f(x) = f_1(x) + i f_2(x)$, $g(x) = g_1(x) + i g_2(x)$, 其中 $f_k(x), g_k(x)$ 为博雷尔函数, $k = 1, 2$. 证明, 如果 $\mathbf{E}|f(\xi)| < \infty$, $\mathbf{E}|g(\xi)| < \infty$, 则有

$$\mathbf{E}|f(\xi) g(\eta)| < \infty$$

和

$$\mathbf{E} f(\xi) g(\eta) = \mathbf{E} f(\xi) \cdot \mathbf{E} g(\eta).$$

(请记住, 我们有: $\mathbf{E} f(\xi) = \mathbf{E} f_1(\xi) + i \mathbf{E} f_2(\xi)$, $\mathbf{E}|f(\xi)| = \mathbf{E}(f_1^2(\xi) + f_2^2(\xi))^{1/2}$.)

2. 设 $\xi = (\xi_1, \cdots, \xi_n)$, 有 $\mathbf{E}\|\xi\|^n < \infty$, 其中 $\|\xi\| = \sqrt{\sum \xi_i^2}$. 证明

$$\varphi_\xi(t) = \sum_{k=0}^n \frac{i^k}{k!} \mathbf{E}(t, \xi)^k + \varepsilon_n(t) \|t\|^n,$$

其中 $t = (t_1, \cdots, t_n)$, $(t, \xi) = t_1 \xi_1 + \cdots + t_n \xi_n$, 而当 $\|t\| \to 0$ 时, 有 $\varepsilon_n(t) \to 0$.

提示: 与一维场合下的证明相类似, 只是需将 $t\xi$ 换为 (t, ξ).

3. 试对 n 元分布函数 $F = F(x_1, \cdots, x_n)$ 与 $G = G(x_1, \cdots, x_n)$ 证明定理 2.

4. 设 $F = F(x_1, \cdots, x_n)$ 为 n 元分布函数, $\varphi = \varphi(t_1, \cdots, t_n)$ 是它所对应的特征函数. 试运用第 3 节 (12) 式中的记号, 证明多元*逆转公式* (*反演公式*) 的正确性:

$$\mathbf{P}(a, b] = \lim_{c \to \infty} \frac{1}{(2\pi)^n} \int_{-c}^{c} \cdots \int_{-c}^{c} \prod_{k=1}^n \frac{e^{-it_k a_k} - e^{-it_k b_k}}{it_k} \varphi(t_1, \cdots, t_n) dt_1 \cdots dt_n.$$

(在上述公式中, 假定区间 $(a, b]$ 是函数 $\mathbf{P}(a, b]$ 的连续区间, 其中 $a = (a_1, \cdots, a_n)$, $b = (b_1, \cdots, b_n)$, 意即对所有 $k = 1, \cdots, n$, 点 a_k, b_k 都是边缘分布函数 $F_k(x_k)$ 的连续点, 其中, $F_k(x_k)$ 是通过在 $F(x_1, \cdots, x_n)$ 中令除了 x_k 以外的所有其他变元趋于无穷时的极限.)

5. 设 $\varphi_k(t)$, $k \geqslant 1$, 都是特征函数, λ_k, $k \geqslant 1$, 均为非负实数, 有 $\sum \lambda_k = 1$. 证明, $\sum \lambda_k \varphi_k(t)$ 是特征函数.

6. 如果 $\varphi(t)$ 是特征函数, 那么 $\operatorname{Re} \varphi(t)$ 与 $\operatorname{Im} \varphi(t)$ 是否也是特征函数?

提示: 设 $\varphi = \varphi(t)$ 是分布 P 的特征函数. 为了回答第一个 (关于 $\operatorname{Re} \varphi(t)$ 的) 问题, 可以考察分布 Q, 其定义为 $Q(A) = \frac{1}{2}[P(A) + P(-A)]$, 其中 $-A = \{-x : x \in A\}$. 而为了回答第二个问题, 可考察函数 $\varphi(t) \equiv 1$.

7. 设 $\varphi_1, \varphi_2, \varphi_3$ 都是特征函数, 现知 $\varphi_1 \varphi_2 = \varphi_1 \varphi_3$. 试问, 是否必有 $\varphi_2 = \varphi_3$?

8. 试证明表 4 与表 5 中所引入的关于特征函数的公式的正确性.

提示: 对于前 5 个离散分布, 它们的特征函数可以通过初等计算得到.

对于负二项分布 ($C_{k-1}^{r-1} p^r q^{k-r}$, $k = r, r+1, \cdots$; $r = 1, 2, \cdots$), 注意到下列事实是有益的: 对于 $|z| < 1$, 有

$$\sum_{k=r}^\infty C_{k-1}^{r-1} z^{k-r} = (1 - z)^{-r}.$$

在求正态分布 $\mathcal{N}(m, \sigma^2)$ 的特征函数 $\varphi(t)$ 时, 需要指出: 根据复变函数理论(在 $m = 0$, $\sigma^2 = 1$ 的场合下), 有

$$\varphi(t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{itx} e^{-x^2/2} dx = e^{t^2/2} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-(x-it)^2/2} dx = e^{t^2/2} \int_L f(z) dz,$$

其中 $f(z) = \frac{1}{\sqrt{2\pi}} e^{-(z-it)^2/2}$, $L = \{z : \operatorname{Im} z = 0\}$, 而

$$\int_L f(z) dz = \int_{L'} f(z) dz = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-y^2/2} dy = 1,$$

其中 $L' = \{z:\ \operatorname{Im} z = t\}$.

类似的处理办法也可用于伽玛分布.

对于计算柯西分布的特征函数 $\varphi(t)$, 运用残数理论是最方便的: 如果 $t > 0$, 则

$$\varphi(t) = \int_{\mathbf{R}} \frac{e^{itx}\theta}{\pi(x^2+\theta^2)} dx = \int_L f(z)dz,$$

其中 $L = \{z:\ \operatorname{Im} z = 0\}$, 而 $f(z) = \frac{e^{itz}}{\pi(z^2+\theta^2)}$. 根据关于残数的柯西定理和若尔当引理 (参阅 [79] 第一卷), 知

$$\int_L f(z)dz = 2\pi i \operatorname*{res}_{i\theta} f = e^{-t\theta}.$$

同理, 对 $t < 0$, 有 $\varphi(t) = e^{t\theta}$. 总之, 我们有 $\varphi(t) = e^{-\theta|t|}$.

9. 设 ξ 为整值随机变量, $\varphi_\xi(t)$ 是其特征函数. 证明,

$$\mathbf{P}\{\xi = k\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ikt}\varphi_\xi(t)dt, \quad k = 0, \pm1, \pm2, \cdots.$$

10. 证明, 在赋有勒贝格测度的空间 $L^2 = L^2([-\pi,\pi], \mathscr{B}[-\pi,\pi])$ 中, 函数族

$$\left\{\frac{1}{\sqrt{2\pi}} e^{i\lambda n},\ n = 0, \pm1, \pm2, \cdots\right\}$$

形成一组正交基.

提示: 可按照如下思路证明之:

(a) 根据 $\varepsilon > 0$, 找到 $c > 0$, 使得

$$\|\varphi - f\|_{L^2} < \varepsilon,$$

其中 $f(x) = \varphi(x) I(|\varphi(x)| \leqslant c)$.

(b) 根据鲁金定理 (参阅第 10 节第 37 题), 找到连续函数 $f_\varepsilon(x)$, 有 $|f_\varepsilon(x)| \leqslant c$, 使得

$$\mu\{x \in [-\pi,\pi]:\ f_\varepsilon(x) \neq f(x)\} < \varepsilon,$$

此时有 $\|f - f_\varepsilon\|_{L^2} \leqslant 2c\sqrt{\varepsilon}$.

(c) 找连续函数 $\widetilde{f}_\varepsilon(x)$, 使得 $\widetilde{f}_\varepsilon(-\pi) = \widetilde{f}_\varepsilon(\pi)$ 和 $\|\widetilde{f}_\varepsilon - f_\varepsilon\|_{L^2} \leqslant \varepsilon$.

(d) 根据魏尔斯特拉斯定理, 找函数 $\overline{f}_\varepsilon(x) = \sum_{k=-n}^{n} a_k e^{ikx}$, 使得

$$\sup_{x\in[-\pi,\pi]} |\widetilde{f}_\varepsilon(x) - \overline{f}_\varepsilon(x)| \leqslant \varepsilon,$$

这也就意味着 $\|\widetilde{f}_\varepsilon - \overline{f}_\varepsilon\|_{L^2} \leqslant \varepsilon$.

由 (a) ~ (d) 推知, 形如 $\sum_{k=-n}^{n} a_k e^{ikx}$ 的有限和在 L^2 中稠密, 亦即函数族 $\left\{\frac{1}{\sqrt{2\pi}} e^{i\lambda n},\ n = 0, \pm1, \pm2, \cdots\right\}$ 形成一组正交基.

11. 在博赫纳–辛钦定理中假定所考察的函数 $\varphi(x)$ 是连续的. 试证明如下的 (里斯的) 结果, 它展示了可以在多大的程度上拒绝关于连续函数的假设.

设 $\varphi = \varphi(t)$ 是勒贝格可测的复值函数, 有 $\varphi(0) = 1$. 那么 $\varphi = \varphi(t)$ 是正定的, 当且仅当, 它 (关于数轴上的勒贝格测度几乎处处) 重合于一个特征函数.

12. 如下函数中, 哪些是特征函数?

$$\varphi(t) = e^{-|t|^\alpha},\quad 0 \leqslant \alpha \leqslant 2, \qquad \varphi(t) = e^{-|t|^\alpha},\quad \alpha > 2,$$

$$\varphi(t) = (1+|t|)^{-1}, \qquad \varphi(t) = (1+t^4)^{-1},$$

$$\varphi(t) = \begin{cases} 1 - |t|^3, & |t| \leqslant 1, \\ 0, & |t| > 1, \end{cases} \qquad \varphi(t) = \begin{cases} 1 - |t|, & |t| \leqslant 1/2, \\ 1/(4|t|), & |t| > 1/2. \end{cases}$$

提示: 为验证其中的一些函数不是特征函数, 最好是回到定理 1, 或者利用下面的第 21 题所介绍的不等式.

13. 证明如下的函数不是特征函数:

$$\varphi(t) = \begin{cases} \sqrt{1-t^2}, & |t| \leqslant 1, \\ 0, & |t| > 1. \end{cases}$$

试问, 函数 $\varphi(t) = \frac{\sin t}{t}$ 是否为特征函数?

14. 证明, 如果函数 $\varphi(t)$ 是特征函数, 则函数 $|\varphi(t)|^2$ 也是特征函数.

15. 证明, 如果函数 $\varphi = \varphi(t)$ 是特征函数, 则对任何 $\lambda \geqslant 0$, 函数 $e^{\lambda(\varphi(t)-1)}$ 都是特征函数. 试问, 函数 $\varphi(t) = e^{\lambda(e^{-|t|}-1)}$ 是否为特征函数?

16. 设函数 $\varphi(t)$ 是特征函数. 证明, 如下的函数都是特征函数:

$$\int_0^1 \varphi(ut)du, \quad \int_0^\infty e^{-u}\varphi(ut)du.$$

17. 证明, 对任何 $n \geqslant 1$, 如下的函数都是特征函数:

$$\varphi_n(t) = \frac{e^{it} - \sum_{k=0}^{n-1} (it)^k/k!}{(it)^n/n!}.$$

18. 设随机变量 X_n 服从区间 $(-n, n)$ 上的均匀分布, 将其特征函数记为 $\varphi_{X_n}(t)$. 证明,

$$\lim_{n\to\infty} \varphi_{X_n}(t) = \begin{cases} 1, & t = 0, \\ 0, & t \neq 0. \end{cases}$$

19. 设 X 为随机变量, $F = F(x)$ 是其分布函数, 而 $(m^{(n)})_{n\geqslant1}$ 是其矩序列

$$\left(m^{(n)} = \int_{-\infty}^{\infty} x^n dF(x)\right).$$

证明, 如果对某个 $s > 0$, 有级数 $\sum_{n=1}^{\infty} \frac{m^{(n)}}{n!} s^n$ 收敛, 则分布函数 $F = F(x)$ 由矩序列 $(m^{(n)})_{n \geqslant 1}$ 唯一决定.

20. 设 $\varphi(t) = \int_{-\infty}^{\infty} e^{itx} dF(x)$ 是分布 $F = F(x)$ 的特征函数. 证明,

$$\lim_{c\to\infty} \frac{1}{2c} \int_{-c}^{c} e^{-itx} \varphi(t) dt = F(x) - F(x-),$$

$$\lim_{c\to\infty} \frac{1}{2c} \int_{-c}^{c} |\varphi(t)|^2 dt = \sum_{x\in\mathbb{R}} [F(x) - F(x-)]^2.$$

特别地, 分布函数 $F = F(x)$ 连续, 当且仅当, 其特征函数 $\varphi(t)$ 满足条件

$$\lim_{c\to\infty} \frac{1}{c} \int_{-c}^{c} |\varphi(t)|^2 dt = 0.$$

21. 证明, 任何特征函数 $\varphi = \varphi(t)$ 都必然满足下列各不等式:

$$1 - \mathrm{Re}\,\varphi(nt) \leqslant n[1 - (\mathrm{Re}\,\varphi(t))^n] \leqslant n^2[1 - \mathrm{Re}\,\varphi(t)], \quad n = 0, 1, 2, \cdots; \quad (*)$$

$$|\mathrm{Im}\,\varphi(t)|^2 \leqslant \frac{1}{2}[1 - \mathrm{Re}\,\varphi(2t)]; \qquad 1 - \mathrm{Re}\,\varphi(2t) \geqslant 2(\mathrm{Re}\,\varphi(t))^2;$$

$$|\varphi(t) - \varphi(s)|^2 \leqslant 4\varphi(0)[1 - \varphi(t-s)]; \qquad 1 - |\varphi(2t)|^2 \leqslant 4[1 - |\varphi(t)|^2];$$

$$|\varphi(t) - \varphi(s)|^2 \leqslant 2[1 - \mathrm{Re}\,\varphi(t-s)];$$

$$\frac{1}{2h} \int_{t-h}^{t+h} \varphi(u) du \leqslant (1 + \mathrm{Re}\,\varphi(h))^{1/2}, \quad t > 0.$$

(最后两个不等式称为拉伊科夫 (Райков) 不等式.)

提示: 上述各不等式的证明全都离不开对公式 $\varphi(t) = \int_{-\infty}^{\infty} e^{itx} dF(x)$ (以及关于 $\mathrm{Re}\,\varphi(t)$ 和 $\mathrm{Im}\,\varphi(t)$ 的相应公式) 的分析. 例如, 为了证明不等式

$$1 - \mathrm{Re}\,\varphi(2t) \leqslant 4[1 - \mathrm{Re}\,\varphi(t)] \qquad (**)$$

(不等式 $(*)$ 的 $n = 2$ 时的特殊情形), 需要用到如下的关系式:

$$1 - \mathrm{Re}\,\varphi(2t) = \int_{-\infty}^{\infty} (1 - \cos 2tx) dF(x) \quad 和 \quad 1 - \cos 2tx \leqslant 4(1 - \cos tx).$$

22. 设 $\varphi(t)$ 为特征函数, 有 $\varphi(t) = 1 + f(t) + o(t^2)$, $t \to 0$, 其中 $f(t) = -f(-t)$. 证明 $\varphi(t) \equiv 1$.

提示: 运用上题中的不等式 $(**)$.

23. 设 X 为随机变量, 其特征函数与分布函数分别为 $\varphi = \varphi(t)$ 和 $F = F(x)$.

(a) 证明, $\int_{-\infty}^{\infty} |x| dF(x) < \infty$, 当且仅当, $\int_{-\infty}^{\infty} \frac{1-\mathrm{Re}\,\varphi(t)}{t^2} dt < \infty$. 并且在该条件下, 有

$$\mathbf{E}|X| \equiv \int_{-\infty}^{\infty} |x| dF(x) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1 - \mathrm{Re}\,\varphi(t)}{t^2} dt = \frac{2}{\pi} \int_{0}^{\infty} \frac{1 - \mathrm{Re}\,\varphi(t)}{t^2} dt.$$

(b) 证明, 如果 $\int_{-\infty}^{\infty} |x| dF(x) < \infty$, 则

$$\mathbf{E}|X| \equiv \int_{-\infty}^{\infty} |x| dF(x) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\mathrm{Re}\,\varphi'(t)}{t} dt.$$

(亦可参阅第 37 题.)

提示: 为证 (a), 只需利用

$$|x| = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1 - \cos xt}{t^2} dt.$$

为证 (b), 需要利用 $|x| = x \,\mathrm{sign}\, x$, 其中

$$\mathrm{sign}\, x = \begin{cases} 1, & x > 0, \\ 0, & x = 0, \\ -1, & x < 0 \end{cases}$$

和

$$\mathrm{sign}\, x = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sin xt}{t} dx.$$

24. 设 $\varphi(t)$ 为特征函数, 有 $\varphi(t) = 1 + O(|t|^\alpha)$, $t \to 0$, 其中 $\alpha \in (0, 2]$. 证明, 对于以 $\varphi(t)$ 作为自己的特征函数的随机变量 ξ, 有

$$\mathbf{P}\{|\xi| > x\} = O(x^{-\alpha}), \quad x \to 0.$$

25. 设随机变量 X 与 Y 独立同分布, 均值为 0, 方差为 1. 试通过考察特征函数, 证明, 如果 $(X+Y)/\sqrt{2}$ 的分布与 X 和 Y 的分布 F 相同, 则 F 是正态分布.

26. 以 $F = F(x)$ 为分布函数的非负随机变量的 X 的拉普拉斯变换 $\widehat{F} = \widehat{F}(\lambda)$ 的定义是: 对 $\lambda \geqslant 0$, 有

$$\widehat{F}(\lambda) = \mathbf{E} e^{-\lambda X} = \int_{[0,\infty)} e^{-\lambda x} dF(x).$$

证明如下的 (伯恩斯坦) 准则: $(0, \infty)$ 上的函数 $\widehat{F} = \widehat{F}(\lambda)$ 是 $[0, \infty)$ 上的某个分布函数 $F = F(x)$ 的拉普拉斯变换, 当且仅当, 该函数是完全单调的(意即对一切 $n \geqslant 0$, 都存在导数 $\widehat{F}^{(n)}(\lambda)$, 并且 $(-1)^n \widehat{F}^{(n)}(\lambda) \geqslant 0$).

27. 设分布函数 $F=F(x)$ 具有密度函数 $f=f(x)$ 和特征函数 $\varphi=\varphi(t)$. 假设有如下二条件之一成立:

$$\text{(a)}\quad \int_{-\infty}^{\infty}|\varphi(t)|dt<\infty \quad \text{或} \quad \text{(b)}\quad \int_{-\infty}^{\infty}f^2(x)dx<\infty.$$

证明, 此时成立如下的帕塞瓦尔公式:

$$\int_{-\infty}^{\infty}f^2(x)dx=\frac{1}{2\pi}\int_{-\infty}^{\infty}|\varphi(t)|^2dt \quad (<\infty).$$

(试比较第 11 节的帕塞瓦尔等式 (14).)

28. 设 $\varphi=\varphi(t)$ 是分布函数 $F=F(x)$ 的特征函数. 证明, 如果分布函数 $F=F(x)$ 具有密度函数 $f=f(x)$, 则当 $t\to\infty$ 时, 有 $\varphi(t)\to 0$.

29. 设 $F=F(x)$ 与 $\widetilde{F}=\widetilde{F}(x)$ 是 $(\mathbb{R},\mathscr{B}(\mathbb{R}))$ 上的两个分布函数, 它们的特征函数分别为 $\varphi=\varphi(t)$ 和 $\widetilde{\varphi}=\widetilde{\varphi}(t)$. 证明如下的帕塞瓦尔等式: 对 $t\in\mathbb{R}$, 有

$$\int\widetilde{\varphi}(x-t)dF(x)=\int e^{-ity}\varphi(y)d\widetilde{F}(y). \tag{*}$$

如果 $\widetilde{F}$ 是正态 $\mathcal{N}(0,\sigma^2)$ 的分布函数, 则

$$\int_{-\infty}^{\infty}e^{-\frac{\sigma^2(x-t)^2}{2}}dF(x)=\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}e^{-ity}e^{-\frac{y^2}{2\sigma^2}}\varphi(y)dy. \tag{**}$$

(试比较第 40 题.)

30. 试由上题中的 (**) 式推出如下结果: 如果分布函数 F_1 与 F_2 的特征函数相同, 则有 $F_1=F_2$. (试比较第 41 题.)

31. 试由第 29 题中的 (*) 式推出如下结果: 如果 $\varphi_\xi(t)$ 是随机变量 ξ 的特征函数, 则随机变量 $|\xi|$ 的拉普拉斯变换是

$$\mathbf{E}e^{-\lambda|\xi|}=\int_{-\infty}^{\infty}\frac{\lambda}{\pi(\lambda^2+t^2)}\varphi_\xi(t)dt,\quad \lambda>0.$$

(参阅第 23 题中的断言.)

32. 设 $F=F(x)$ 为分布函数, 而 $\varphi(t)=\int_{-\infty}^{\infty}e^{itx}dF(x)$ 是它的特征函数. 试根据定理 3 中的断言 b), 证明, 性质 $\int_{-\infty}^{\infty}|\varphi(t)|dt<\infty$ 可以保证该分布具有连续的密度函数 $f(x)$.

试举例说明, 存在分布函数 $F=F(x)$, 它具有连续的密度函数, 但却有 $\int_{-\infty}^{\infty}|\varphi(t)|dt=\infty$.

33. 设 $\varphi(t)=\int_{\mathbb{R}}e^{itx}dF(x)$ 是分布函数 $F(x)$ 的特征函数. 如所周知 (定理 1), 如果 $\int_{\mathbb{R}}|x|dF(x)<\infty$, 则 $\varphi(t)$ 可微. 试举例说明, 一般来说, 其逆不真. 甚至存在这样的特征函数 $\varphi(t)$, 它无限次可微, 但是却有 $\int_{\mathbb{R}}|x|dF(x)=\infty$. 试给出一个这样的特征函数的例子.

34. (关于逆转公式) 仿照定理 3 的证明, 证明, 对于任何分布函数 $F=F(x)$ 与任何 $a<b$, 都有如下的 (广义) "逆转公式" 成立:

$$\lim_{c\to\infty}\frac{1}{2\pi}\int_{-c}^{c}\frac{e^{-ita}-e^{-itb}}{it}\varphi(t)dt=\frac{1}{2}[F(b)+F(b-)]-\frac{1}{2}[F(a)+F(a-)].$$

35. (a) 具有密度函数

$$f(x)=\frac{1-\cos x}{\pi x^2},\quad x\in\mathbb{R}$$

的概率分布的特征函数是:

$$\varphi(t)=\begin{cases}1-|t|, & |t|\leqslant 1,\\ 0, & |t|>1.\end{cases}$$

(b) 试问, 具有如下的密度函数的概率分布的特征函数是怎样的?

$$f(x)=\frac{1-\cos\pi x}{\pi^2x^2},\quad x\in\mathbb{R}.$$

(c) 证明, 分别具有下列密度函数

$$f_1(x)=\frac{1}{\pi\operatorname{ch}x},\quad f_2(x)=\frac{1}{2\operatorname{ch}^2x},\quad x\in\mathbb{R}$$

的概率分布的特征函数分别为:

$$\varphi_1(t)=\frac{1}{\operatorname{ch}\frac{1}{2}\pi t},\quad \varphi_2(t)=\frac{\pi t}{2\operatorname{sh}\frac{1}{2}\pi t}.$$

(d) 试求与下列各个特征函数相对应的概率分布:

$$\frac{1+it}{1+t^2},\quad \frac{1-it}{1+t^2},\quad \cos\frac{t}{2},\quad \frac{2}{3e^{it}-1},\quad \frac{1}{2}e^{-it}+\frac{1}{3}+\frac{1}{6}e^{2it}.$$

36. 设 $m^{(k)}=\int_{\mathbb{R}}x^kdF(x)\,(k\geqslant 1)$ 是分布 $F=F(x)$ 的实质有限矩. 证明

$$\int_{\mathbb{R}}\operatorname{ch}(ax)dF(x)=\sum_k\frac{a^{2k}}{(2k)!}m^{(2k)},$$

其中 $\operatorname{ch}y=(e^{-y}+e^y)/2$.

37. 如同第 23 题, 设以 $F=F(x)$ 为分布函数的随机变量 X 的特征函数是 $\varphi=\varphi(t)$. 证明, 对 $\beta\in(0,2)$, 有

$$\int_{-\infty}^{\infty}|x|^\beta dF(x)<\infty$$

当且仅当

$$\int_{-\infty}^{\infty}\frac{1-\operatorname{Re}\varphi(t)}{|t|^{1+\beta}}dt<\infty;$$

并且此时有

$$\mathbf{E}|X|^\beta = \int_{-\infty}^{\infty} |x|^\beta dF(x) = C_\beta \int_{-\infty}^{\infty} \frac{1-\operatorname{Re}\varphi(t)}{|t|^{1+\beta}}dt,$$

其中

$$C_\beta = \left[\int_{-\infty}^{\infty} \frac{1-\cos t}{|t|^{1+\beta}}dt\right]^{-1} = \frac{\Gamma(1+\beta)}{\pi}\sin\frac{\beta\pi}{2}.$$

提示: 利用关系式

$$|x|^\beta = C_\beta \int_{-\infty}^{\infty} \frac{1-\cos xt}{|t|^{1+\beta}}dt.$$

38. 在第 8 节第 27 题中, 通过直接计算随机变量 ξ/η, $\xi/|\eta|$, $|\xi|/|\eta|$, 以及服从以 $\frac{1}{\pi(1+x^2)}$ $(x\in\mathbb{R})$ 为密度函数的柯西分布的随机变量 C 的特征函数来证明该题中的断言.

39. 设 $\xi_1,\xi_2,\cdots$ 为独立同分布的标准正态 $\mathcal{N}(0,1)$ 随机变量. 对 $n\geqslant 1$, 令

$$X_n = \xi_1+\cdots+\xi_n$$

和

$$S_n = X_1+\cdots+X_n.$$

试运用特征函数方法, 证明

$$\frac{S_n}{n^{3/2}} \xrightarrow{\text{law}} \mathcal{N}(0,\sigma^2),$$

其中 $\mathcal{N}(0,\sigma^2)$ 为正态分布, $\sigma^2=1/3$.

40. 设 ξ 与 η 为相互独立的随机变量, 其中 η 服从标准正态分布 $(\eta\sim\mathcal{N}(0,1))$. 又设 $f=f(x)$ 为具有紧支撑的有界博雷尔函数. 证明, 对 $\sigma>0$, 有

$$\mathbf{E}f\left(\xi+\frac{1}{\sigma}\eta\right) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{t^2}{2\sigma^2}}\varphi_\xi(t)\widehat{f}(-t)dt, \qquad (*)$$

其中, $\varphi_\xi(t)=\mathbf{E}e^{it\xi}$, 而 $\widehat{f}(t)=\int_{-\infty}^{\infty} e^{it\xi}f(x)dx$ (试比较第 29 题).

试对随机向量 ξ 与 η 陈述相应的结论.

41. 试利用上题中的 $(*)$ 式, 证明, 随机变量 ξ 的特征函数 $\varphi_\xi(t)$ 可以唯一确定其分布. (试比较定理 2.)

提示: 试证明, 在上题中的条件下, 可以由 $(*)$ 式推出

$$\mathbf{E}f(\xi) = \lim_{\sigma\to\infty}\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{t^2}{2\sigma^2}}\varphi_\xi(t)\widehat{f}(-t)dt \qquad (**)$$

(利用具有紧支撑的有界博雷尔函数 $f=f(x)$ 的任意性), 即可由此断言: 特征函数 $\varphi_\xi(t)$ 完全可以确定随机变量 ξ 的分布. ($(**)$ 式对于多元随机向量也成立. 试由多元场合下的 $(*)$ 式推出相应的 $(**)$ 式.)

42. 设 $\varphi=\varphi(t)$ 为特征函数, 对某 $b>0$, 有

$$\varphi(t)\leqslant a, \quad 对于 \quad |t|\geqslant b,$$

其中 $0<a<1$. 证明, 在 $|t|<b$ 时, 有如下的克拉默不等式成立:

$$\varphi(t)\leqslant 1-(1-a^2)\frac{t^2}{8b^2}.$$

提示: 利用第 21 题中的不等式:

$$1-|\varphi(2t)|^2\leqslant 4[1-|\varphi(t)|^2].$$

43. (补充第 21 题中的不等式) 设 $\varphi=\varphi(t)$ 是分布 $F=F(x)$ 的特征函数. 证明如下的贝里–埃森不等式: 对 $1\leqslant r\leqslant 2$, 有

$$|1-\varphi(t)|\leqslant C_r\beta_r|t|^r,$$

其中 $\beta_r=\int_{-\infty}^{\infty}|x|^r dF(x)$, 而 C_r 为常数.

44. 对于整数 $n\geqslant 1$, 随机变量 X 的 n 阶矩 $m^{(n)}=\mathbf{E}X^n$ 和 n 阶绝对矩 $\beta_n=\mathbf{E}|X|^n$ 都可以有直接的表达式, 其中出现有 X 的特征函数 $\varphi(t)=\mathbf{E}e^{itX}$ $(t\in\mathbb{R})$ 的 n 阶导数 (参阅, 例如, (13) 式或第 23 题). 对于非整数的 $\alpha>0$, 为了得到关于 $m^{(\alpha)}=\mathbf{E}X^\alpha$ 和 $\beta_\alpha=\mathbf{E}|X|^\alpha$ 的相应结果, 需要用到按照如下方式引入的分数阶导数.

写 $\alpha=n+a$, 其中 n 为非负整数, 而 $0<a<1$. 我们将如下的函数称为函数 $f=f(t)\,(t\in\mathbb{R})$ 的 α 阶分数阶导数 (假定相应的积分存在):

$$\frac{a}{\Gamma(1-a)}\int_{-\infty}^{t}\frac{f^{(n)}(t)-f^{(n)}(s)}{(t-s)^{1+\alpha}}ds,$$

并记为 $D^{(\alpha)}f(t)\ (=\frac{d^\alpha}{dt^\alpha}f(t))$. 特别地, 如果 $f(t)=\varphi(t)(=\int_{-\infty}^{\infty}e^{itx}dF(x))$ 是特征函数, 则

$$D^{(n+a)}\varphi(t)\Big|_{t=0} = -\frac{1}{\Gamma(-a)}\int_0^\infty \frac{f^{(n)}(0)-f^{(n)}(-u)}{u^{1+a}}du$$

$$= \frac{(-1)^{n+1}}{\Gamma(-a)}\left\{\int_0^\infty\left[\int_{-\infty}^{\infty}x^n(1-\cos ux)dF(x)+i\int_{-\infty}^{\infty}x^n\sin ux dF(x)\right]\frac{du}{u^{1+a}}\right\}. \qquad (*)$$

证明, 对于偶数 n, 绝对矩 $\beta_{n+a}<\infty$, 当且仅当

(i) $\beta_n<\infty$,

(ii) $\operatorname{Re}\left[D^{(n+a)}\varphi(t)\Big|_{t=0}\right]$ 存在.

并且在此种情况下, 有

$$\beta_{n+a} = \frac{1}{\cos\frac{a\pi}{2}}\operatorname{Re}\left[(-1)^{n/2}D^{(\alpha)}\varphi(t)\Big|_{t=0}\right].$$

提示: 为了证明结论, 需要利用 (*) 式, 以及对于所有 $0<b<2$ 都成立的如下公式:

$$\int_0^\infty \frac{1-\cos u}{u^{1+b}}du = -\Gamma(-b)\cos\frac{\pi b}{2}.$$

注: 对于任何 $n \geqslant 0$ 和 $0<a<1$, 关于 $\mathbf{E}|X|^{n+a}$ 的更加详细的讨论见专著 [77].

45. 设 $\varphi=\varphi(t)$ 为特征函数. 证明, 对任何 u 与 s, 都有

$$|\varphi(u+s)| \geqslant |\varphi(u)|\cdot|\varphi(s)| - [1-|\varphi(u)|^2]^{1/2}[1-|\varphi(s)|^2]^{1/2}.$$

提示: 利用博纳赫–辛钦定理 (第 6 小节).

46. 设随机变量对 (ξ,η) 具有密度函数

$$f(x,y)=\frac{1}{4}\{1+xy(x^2-y^2)\}I(|x|<1,\,|y|<1).$$

证明, 随机变量 ξ 与 η 相互独立, 并且分别具有密度函数:

$$f_\xi(x)=\frac{1}{2}I(|x|<1), \quad f_\eta(y)=\frac{1}{2}I(|y|<1).$$

而且 $\xi+\eta$ 的特征函数 $\varphi_{\xi+\eta}(t)$ 等于特征函数 $\varphi_\xi(t)$ 与 $\varphi_\eta(t)$ 的乘积, 亦即 $\varphi_{\xi+\eta}(t)=\varphi_\xi(t)\cdot\varphi_\eta(t)$.

47. 设 $\xi_1,\xi_2,\cdots$ 为独立同分布的随机变量序列, 分别以 $1/10$ 的概率取值 $0,1,\cdots,9$. 令

$$X_n=\sum_{k=1}^n \frac{\xi_k}{10^k}.$$

证明, $(X_n)_{n\geqslant 1}$ 不仅依分布, 而且几乎必然地收敛到一个服从区间 $[0,1]$ 上的均匀分布的随机变量.

提示: 利用特征函数方法.

§13. 高斯系

1. 证明, 在高斯系 $(\xi,\eta_1,\cdots,\eta_n)$ 场合下, 条件数学期望 $\mathbf{E}(\xi|\eta_1,\cdots,\eta_n)$ 等于广义数学期望 $\widehat{\mathbf{E}}(\xi|\eta_1,\cdots,\eta_n)$.

2. 设 $(\xi,\eta_1,\cdots,\eta_k)$ 为高斯系. 试说明条件数学期望 $\mathbf{E}(\xi^n|\eta_1,\cdots,\eta_k)$ 的结构, $n\geqslant 1$ (作为 $\eta_1,\cdots,\eta_k$ 的函数).

3. 设 $X=(X_k)_{1\leqslant k\leqslant n}$ 与 $Y=(Y_k)_{1\leqslant k\leqslant n}$ 是两个高斯序列, 有 $\mathbf{E}X_k=\mathbf{E}Y_k$, $\mathbf{D}X_k=\mathbf{D}Y_k$, $1\leqslant k\leqslant n$ 及

$$\mathbf{cov}\,(X_k,X_l)\leqslant \mathbf{cov}\,(Y_k,Y_l), \quad 1\leqslant k,\ l\leqslant n.$$

证明如下的斯莱皮恩 (Slepian) 不等式: 对 $x\in\mathbb{R}$, 有

$$\mathbf{P}\left\{\sup_{1\leqslant k\leqslant n} X_k<x\right\}\leqslant \mathbf{P}\left\{\sup_{1\leqslant k\leqslant n} Y_k<x\right\}.$$

4. 设 ξ_1,ξ_2,ξ_3 为相互独立的高斯随机变量, 有 $\xi_i\sim\mathcal{N}(0,1)$, $i=1,2,3$. 证明,

$$\frac{\xi_1+\xi_2\xi_3}{\sqrt{1+\xi_3^2}}\sim\mathcal{N}(0,1).$$

(由此产生出一个有趣的问题: 即对 $n\geqslant 2$, 考察 n 个相互独立的高斯随机变量 $\xi_1,\cdots,\xi_n$. 在关于它们的怎样的非线性变换之下, 所得的随机变量仍然服从高斯分布?)

5. 设 $r(s,t)$ 是 (25),(29) 和 (30) 式所定义的函数. 证明, 矩阵 $\mathbb{R}=(r(s,t)_{s,t\in\mathfrak{U}})$ 是非负定的.

6. 设 A 是某个 $m\times n$ 矩阵. 称 $n\times m$ 矩阵 $A^\otimes$ 是 A 的*伪逆矩阵*, 如果存在矩阵 U 与 V, 使得

$$AA^\otimes A=A, \quad A^\otimes=UA^*=A^*V.$$

证明, 如此定义的矩阵 $A^\otimes$ 存在且唯一.

7. 证明, 如果在关于正常相关的定理中的公式 (19) 和 (20) 中, 将 $\mathbf{D}_{\xi\xi}^{-1}$ 换成如上所述的广义逆 $\mathbf{D}_{\xi\xi}^\otimes$, 则这两个公式对于退化的矩阵 $\mathbf{D}_{\xi\xi}$ 也成立.

8. 设 $(\theta,\xi)=(\theta_1,\cdots,\theta_k;\xi_1,\cdots,\xi_l)$ 是具有非退化阵 $\Delta\equiv\mathbf{D}_{\theta\theta}-\mathbf{D}_{\xi\xi}^\oplus\mathbf{D}_{\theta\xi}^*$ 的高斯向量. 证明, 分布函数 $\mathbf{P}(\theta\leqslant a|\xi)=\mathbf{P}(\theta_1\leqslant a_1,\cdots,\theta_k\leqslant a_k|\xi)$ 具有由下式所定义的 ($\mathbf{P}$-a.s.) 密度函数 $p(a_1,\cdots,a_k|\xi)$:

$$p(a_1,\cdots,a_k|\xi)=\frac{|\Delta|^{-1/2}}{(2\pi)^{k/2}}\exp\left\{-\frac{1}{2}(a-\mathbf{E}(\theta|\xi))^*\Delta^{-1}(a-\mathbf{E}(\theta|\xi))\right\}.$$

9. 设 ξ 与 η 为相互独立的标准高斯随机变量 (均值为 0, 方差为 1).

(a) 证明, 随机变量 $\xi+\eta$ 与 $\xi-\eta$ 为相互独立的高斯随机变量.

(b) 试利用 (a) 和第 8 节第 27 题中的结论, 证明,

$$C\overset{\text{law}}{=}\frac{\xi+\eta}{\xi-\eta}\overset{\text{law}}{=}\frac{1+\eta/\xi}{1-\eta/\xi}\overset{\text{law}}{=}\frac{1+C}{1-C}\overset{\text{law}}{=}\frac{1}{C},$$

其中随机变量 C 服从以 $\frac{1}{\pi(1+x^2)}$ 为密度的柯西分布 (如同往常, “$\overset{\text{law}}{=}$” 表示同分布).

10. (伯恩斯坦) 设 ξ 与 η 为具有有限方差的独立同分布的随机变量. 证明, 如果随机变量 $\xi+\eta$ 与 $\xi-\eta$ 相互独立, 则 ξ 与 η 是高斯随机变量. (关于该命题的推广, 即 “达尔穆亚–斯基托维奇定理”, 见第 44 题.)

提示: 可以按照如下步骤证明该命题 (以 $\varphi_\xi(t)$ 表示随机变量 ξ 的特征函数):

(a) 由 $\varphi_\xi(t)=\varphi_{\frac{\xi+\eta}{2}}(t)\varphi_{\frac{\xi-\eta}{2}}(t)=\varphi_{\xi+\eta}(t/2)\varphi_{\xi-\eta}(t/2)$ 可以推知

$$\varphi_\xi(t)=\left(\varphi_\xi\left(\frac{t}{2}\right)\right)^2\left|\varphi_\eta\left(\frac{t}{2}\right)\right|^2,\quad t\in\mathbb{R},$$

同理有

$$\varphi_\eta(t)=\left(\varphi_\eta\left(\frac{t}{2}\right)\right)^2\left|\varphi_\xi\left(\frac{t}{2}\right)\right|^2,\quad t\in\mathbb{R}.$$

(b) 由 (a) 推知 $|\varphi_\xi(t)|=|\varphi_\eta(t)|$ 和 $|\varphi_\eta(t)|=\left|\varphi_\xi\left(\frac{t}{2}\right)\right|^4$.

(c) 由 (b) 推出, 对任何 $t\in\mathbb{R}$, 都有 $\varphi_\eta(t)\neq 0$, 因此, 可以定义 $f(t)=\ln|\varphi_\eta(t)|$, 并且对其有 $f(t)=4f(t/2)$, $t\in\mathbb{R}$.

(d) 由 $\mathbf{E}\eta^2<\infty$ 可知 $\varphi_\eta(t)\in C^2(\mathbb{R})$, 再由 (c) 得

$$f''(t)=f''\left(\frac{t}{2}\right)=\cdots=f''\left(\frac{t}{2^k}\right)\to f''(0),\quad t\in\mathbb{R},$$

故知 $f''(t)$ 为常数.

(e) 由 (d) 推得 $f(t)=at^2+bt+c$, 结合 (c), 可知 $f(t)=at^2$.

(f) 由 (e) 推得 $\varphi_\eta(t)=e^{i\alpha(t)+at^2}$, 其中 $\alpha(t)$ 可以取为连续函数, 因为函数 $\varphi_\eta(t)$ 是连续的.

(g) 证明 $\alpha(t)$ 满足性质:

$$\alpha(t)=2\alpha(t/2),\quad t\in\mathbb{R}.$$

(h) 由 $\mathbf{E}\eta^2<\infty$ 推得 $\varphi_\eta(t)$ 在 $t=0$ 处可微, 再由 (g) 推知, 当 $k\to\infty$ 时, 有

$$\frac{\alpha(t)}{t}=\frac{\alpha(t/2^k)}{t/2^k}\to\alpha'(0),\quad t\neq 0,$$

从而有 $\alpha(t)=\alpha'(0)t$.

总而言之, 我们得知 $\varphi_\eta(t)=e^{i\alpha'(0)t+at^2}$, 亦即 η 为高斯随机变量. 同理, ξ 是高斯随机变量.

11. (默瑟定理) 设 $r=r(s,t)$ 是 $[a,b]\times[a,b]$ 上的连续的协方差函数, 其中 $-\infty<a<b<\infty$. 证明, 方程

$$\lambda\int_a^b r(s,t)u(t)dt=u(s),\quad a\leqslant s\leqslant b$$

对于无穷多个数值 $\lambda=\lambda_k<0$, $k\geqslant 1$, 都有着相应的连续解 $u_k=u_k(x)$, $k\geqslant 1$, 这些解构成 $L^2(a,b)$ 中的完备的正交系, 具有性质

$$r(s,t)=\sum_{k=1}^{\infty}\frac{u_k(s)u_k(t)}{\lambda_k},$$

其中的级数在 $[a,b]\times[a,b]$ 中一致绝对收敛.

12. 设 $X=\{X_t,t\geqslant 0\}$ 为高斯过程, 具有 $\mathbf{E}X_t=0$ 和协方差函数 $r(s,t)=e^{-|t-s|}$, $t,s\geqslant 0$. 设 $0<t_1<\cdots<t_n$, 而 $f_{t_1,\cdots,t_n}(x_1,\cdots,x_n)$ 是变量 $X_{t_1},\cdots,X_{t_n}$ 的密度函数. 证明,

$$f_{t_1,\cdots,t_n}(x_1,\cdots,x_n)$$
$$=\left[(2\pi)^n\prod_{j=2}^{n}(1-e^{2(t_{j-1}-t_j)})\right]^{-1/2}\exp\left\{-\frac{x_1^2}{2}-\frac{1}{2}\sum_{j=2}^{n}\frac{(x_j-e^{(t_{j-1}-t_j)}x_{j-1})^2}{1-e^{2(t_{j-1}-t_j)}}\right\}.$$

13. 设 $f=\{f_n=f_n(u),\ n\geqslant 1;\ u\in[0,1]\}$ 是 L^2 函数 (依 $[0,1]$ 上的勒贝格测度) 的完备正交系, 而 $\xi_1,\xi_2,\cdots$ 是独立同分布的 $\mathcal{N}(0,1)$ 随机变量. 证明, 过程 $B_t=\sum\limits_{n\geqslant 1}\xi_n\int_0^t f_n(u)du$, $0\leqslant t\leqslant 1$ 为布朗运动.

14. 证明, 如果 $B^\circ=(B_t^\circ)_{0\leqslant t\leqslant 1}$ 是布朗桥, 则对 $B_t=(1+t)B^\circ_{t/(t+1)}$, 随机过程 $B=(B_t)_{t\geqslant 0}$ 是布朗运动.

15. 设 $B=(B_t)_{t\geqslant 0}$ 是布朗运动. 证明, 如下各过程都是布朗运动:

$$B_t^{(1)}=-B_t,\quad t\geqslant 0;$$
$$B_t^{(2)}=tB_{1/t},\quad t>0,\quad B_0^{(2)}=0;$$
$$B_t^{(3)}=B_{t+s}-B_s,\quad s>0,\quad t\geqslant 0;$$
$$B_t^{(4)}=B_T-B_{T-t},\quad \text{对于 } 0\leqslant t\leqslant T,\quad T>0;$$
$$B_t^{(5)}=\frac{1}{a}B_{a^2t},\quad a>0,\quad t\geqslant 0\quad (\text{自模性}).$$

16. 设 $B^\mu=(B_t+\mu t)_{t\geqslant 0}$ 是带损失的布朗运动.

(a) 试求 $B_{t_1}^\mu+B_{t_2}^\mu$ $(t_1<t_2)$ 的分布.

(b) 试对 $t_0<t_1<t_2$, 求 $\mathbf{E}B_{t_0}^\mu B_{t_1}^\mu$ 和 $\mathbf{E}B_{t_0}^\mu B_{t_1}^\mu B_{t_2}^\mu$.

17. 设 B^μ 是上题中的随机过程. 试分别对 $t_1<t_2$ 和 $t_1>t_2$, 求条件分布

$$\mathbf{P}(B_{t_2}^\mu\in\cdot\mid B_{t_1}^\mu);$$

并对 $t_0<t_1<t_2$, 求条件分布

$$\mathbf{P}(B_{t_2}^\mu\in\cdot\mid B_{t_0}^\mu,\ B_{t_1}^\mu).$$

18. 设 $B=(B_t)_{t\geqslant 0}$ 是布朗运动. 证明, 对于 $Y_t=e^{-t}B_{e^{2t}}$, 过程 $Y=(Y_t)_{t\in\mathbb{R}}$ 是奥恩斯坦–乌伦贝克 (Ornstein-Uhlenbeck) 过程 (即有 $\mathbf{E}Y_t=0$ 与 $\mathbf{E}Y_sY_t=e^{-|t-s|}$ 的高斯–马尔可夫过程).

19. 设 $Y = (Y_t)_{t\in\mathbb{R}}$ 是奥恩斯坦–乌伦贝克过程. 证明,

$$B_t^\circ = \begin{cases} \sqrt{t(1-t)}Y_{\frac{1}{2}\ln\frac{t}{1-t}}, & 0 < t < 1, \\ 0, & t = 0,\ 1 \end{cases}$$

是布朗运动.

20. 设 $\xi_0, \xi_1, \xi_2, \cdots$ 是相互独立的标准高斯 $\mathcal{N}(0,1)$ 随机变量. 证明, 级数

$$B_t^\circ = \sum_{k=1}^{\infty} \xi_k \frac{\sqrt{2}\sin k\pi t}{k\pi}, \quad 0 \leqslant t \leqslant 1$$

事实上定义了一个布朗桥, 而级数

$$B_t = \xi_0 t + \sum_{k=1}^{\infty} \xi_k \frac{\sqrt{2}\sin k\pi t}{k\pi}, \quad 0 \leqslant t \leqslant 1$$

则与级数 (26) 一样, 给出了一个布朗运动.

21. 试详细证明, 由 (26) 与 (28) 式所定义的过程 $(B_t)_{0\leqslant t\leqslant 1}$, 以及过程

$$B_t = \sqrt{2}\sum_{n=1}^{\infty} \xi_n \frac{1-\cos n\pi t}{n\pi}$$

(其中随机变量序列 $(\xi_n)_{n\geqslant 1}$ 与 (26) 和 (28) 式中相同) 都是布朗运动.

22. 设 $X = (X_k)_{1\leqslant k\leqslant n}$ 为高斯序列, 记

$$m = \max_{1\leqslant k\leqslant n} \mathbf{E}X_k, \quad \sigma^2 = \max_{1\leqslant k\leqslant n} \mathbf{D}X_k.$$

今知对某个 $a > 0$, 有

$$\mathbf{P}\left\{\max_{1\leqslant k\leqslant n}(X_k - \mathbf{E}X_k) \geqslant a\right\} \leqslant \frac{1}{2}.$$

证明, 有如下的博雷尔不等式成立:

$$\mathbf{P}\left\{\max_{1\leqslant k\leqslant n} X_k > x\right\} \leqslant 2\Psi\left(\frac{x-m-a}{\sigma}\right),$$

其中 $\Psi(x) = (2\pi)^{-1/2}\int_x^\infty e^{-y^2/2}dy$.

23. 设 (X,Y) 为二维高斯随机变量, 有 $\mathbf{E}X = \mathbf{E}Y = 0$, $\mathbf{E}X^2 > 0$, $\mathbf{E}Y^2 > 0$ 以及相关系数 $\rho = \frac{\mathbf{E}XY}{\sqrt{\mathbf{E}X^2\mathbf{E}Y^2}}$, $|\rho| < 1$.

(a) 证明, 随机变量 X 与 $Z = (Y - \rho X)/\sqrt{1-\rho^2}$ 为相互独立的高斯随机变量.

(b) 证明,

$$\mathbf{P}\{XY < 0\} = 1 - 2\mathbf{P}\{X > 0,\ Y > 0\} = \pi^{-1}\arccos\rho,$$

并由此推出如下各关系式:

$$\mathbf{P}\{X > 0,\ Y > 0\} = \mathbf{P}\{X < 0,\ Y < 0\} = \frac{1}{4} + \frac{1}{2\pi}\arcsin\rho,$$
$$\frac{\partial}{\partial\rho}\mathbf{P}\{X > 0,\ Y > 0\} = \frac{1}{2\pi\sqrt{1-\rho^2}},$$
$$\mathbf{P}\{X > 0,\ Y < 0\} = \mathbf{P}\{X < 0,\ Y > 0\} = \frac{1}{4} - \frac{1}{2\pi}\arcsin\rho.$$

(c) 令 $Z = \max(X, Y)$, 其中 $\mathbf{E}X^2 = \mathbf{E}Y^2 = 1$. 证明

$$\mathbf{E}Z = \sqrt{\frac{1-\rho}{\pi}}, \quad \mathbf{E}Z^2 = 1.$$

(d) 证明, 对任何 a 与 b, 都有:

$$(1-\Phi(a))(1-\Phi(c)) \leqslant \mathbf{P}\{X > a,\ Y > b\} \leqslant (1-\Phi(a))(1-\Phi(c)) + \frac{\rho\varphi(b)(1-\Phi(d))}{\varphi(a)},$$

其中 $c = (b - a\rho)/\sqrt{1-\rho^2}$, $d = (a - b\rho)/\sqrt{1-\rho^2}$, 而 $\varphi(x) = \Phi'(x)$ 为标准正态密度.

提示: 命题 (b) 可以由性质 (a) 推出.

24. 设随机变量 X 与 Y 相互独立, 其中 $X \sim \mathcal{N}(0,1)$, 而 $\mathbf{P}\{Y = 1\} = \mathbf{P}\{Y = -1\} = \frac{1}{2}$. 令 $Z = XY$. 证明, $Z \sim \mathcal{N}(0,1)$, 求随机变量对 (X,Z) 与随机变量对 (Y,Z) 的分布, 并求 $X + Z$ 的分布. 证明, 随机变量 X 与 Z 不相关, 也不独立.

25. 设 ξ 为标准正态随机变量, 即有 $\xi \sim \mathcal{N}(0,1)$. 令

$$\eta_\alpha = \begin{cases} \xi, & \text{如果 } |\xi| \leqslant \alpha, \\ -\xi, & \text{如果 } |\xi| > \alpha. \end{cases}$$

证明, $\eta_\alpha \sim \mathcal{N}(0,1)$, 并且当 α 满足关系式

$$\int_0^\alpha x^2 f_\xi(x)dx = \frac{1}{4} \quad \left(f_\xi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}\right)$$

时, 随机变量 ξ 与 η_α 不相关①, 但却为不相互独立的高斯随机变量 (试比较定理 1 中的断言 a)).

26. 设 ξ 与 η 均为正态分布随机变量, 有 $\mathbf{E}\xi = \mathbf{E}\eta = 0$, $\mathbf{E}\xi^2 = \mathbf{E}\eta^2 = 1$ 和 $\mathbf{E}\xi\eta = \rho$. 证明,

① 原书为: 随机变量 ξ 与 $\eta_{1/4}$ 不相关 —— 译者注.

(a) $\mathbf{E}\max(\xi,\eta)=\sqrt{(1-\rho)/\pi}$.

(b) $\mathbf{E}(\xi\,|\,\eta)=\rho\eta$, $\mathbf{D}(\xi\,|\,\eta)=1-\rho^2$.

(c) $\mathbf{E}(\xi\,|\,\xi+\eta=z)=z/2$, $\mathbf{D}(\xi\,|\,\xi+\eta=z)=(1-\rho)/2$.

(d) $\mathbf{E}(\xi+\eta\,|\,\xi>0,\eta>0)=2\sqrt{2/\pi}$.

试在 $\mathbf{D}\xi=\sigma_1^2$, $\mathbf{D}\eta=\sigma_2^2$ 的场合下考虑相应的问题, 其中 $\sigma_1>0$, $\sigma_2>0$.

27. 设 $\binom{X}{Y}$ 为二维高斯随机变量, 具有相关矩阵

$$\mathbf{cov}\,(X,Y)=\begin{pmatrix}\sigma^2 & \sigma^2\\ \sigma^2 & \sigma^2\end{pmatrix}.$$

试将 $\binom{X}{Y}$ 表示为如下形式:

$$\begin{pmatrix}X\\ Y\end{pmatrix}=Q\begin{pmatrix}\xi\\ \eta\end{pmatrix},$$

其中, Q 为正交矩阵, 而 ξ 与 η 为相互独立的高斯随机变量.

28. 设 $\xi=(\xi_1,\cdots,\xi_n)$ 为具有零均值和相关矩阵 $\mathbb{R}=\|\mathbf{E}\xi_i\xi_j\|$ 的非退化的高斯随机向量. 设 $\mathbb{R}$ 的特征根为 $\lambda_1,\cdots,\lambda_n$. 证明, 随机变量 $\xi_1^2+\cdots+\xi_n^2$ 的特征函数 $\varphi(t)$ 与随机变量 $\lambda_1\eta_1^2+\cdots+\lambda_n\eta_n^2$ 的特征函数相同, 其中 $\eta_1,\cdots,\eta_n$ 为相互独立的标准高斯随机变量 $(\eta_k\sim\mathcal{N}(0,1))$, 并且可将 $\varphi(t)$ 表示为:

$$\varphi(t)=\prod_{j=1}^{n}|1-2it\lambda_j|^{-1/2}.$$

29. 设 $\xi_1,\cdots,\xi_n$ 是独立同分布的随机变量, 其中 $n\geqslant 2$. 证明, 随机向量 $(\xi_1,\cdots,\xi_n)$ 的分布函数为球对称的, 当且仅当, 随机变量 $\xi_1,\cdots,\xi_n$ 中的每一个都服从零均值的正态分布.

提示: 考虑特征函数.

30. (关于正态分布 $\mathcal{N}(m,\sigma^2)$ 的统计量. I.) 设 $\xi_1,\cdots,\xi_n$, $n\geqslant 2$, 为独立同分布的正态 $\mathcal{N}(m,\sigma^2)$ 随机变量. 令

$$\overline{\xi}=\frac{1}{n}\sum_{k=1}^{n}\xi_k,\quad s_1^2=\frac{1}{n-1}\sum_{k=1}^{n}(\xi_k-\overline{\xi})^2.$$

证明, 随机变量 $\overline{\xi}$ 与 s_1^2 相互独立, 并且

$$(n-1)s_1^2\overset{\mathrm{d}}{=}\sum_{k=1}^{n-1}(\xi_k-m)^2.$$

提示: 利用前一题的结论.

31. (关于正态分布 $\mathcal{N}(m,\sigma^2)$ 的统计量. II.) 设 $\xi_1,\cdots,\xi_n$ 为独立同分布的正态 $\mathcal{N}(m,\sigma^2)$ 随机变量, 而 $x=(x_1,\cdots,x_n)$ 是对于 $\xi=(\xi_1,\cdots,\xi_n)$ 的观察值, $n\geqslant 1$.

(a) 证明, 统计量对

$$T_1(x)=\sum_{k=1}^{n}x_k,\quad T_2(x)=\sum_{k=1}^{n}x_k^2$$

与

$$\overline{x}=\frac{1}{n}\sum_{k=1}^{n}x_k,\quad s^2(x)=\frac{1}{n}\sum_{k=1}^{n}(x_k-\overline{x})^2$$

都是充分统计量.

(b) 证明,

$$s^2(x)=\frac{1}{n}\sum_{k=1}^{n}x_k^2-\overline{x}^2.$$

32. (关于正态分布 $\mathcal{N}(m,\sigma^2)$ 的统计量. III: m 未知, $\sigma^2=\sigma_0^2$.) 在本题中, 假设 $\xi_1,\cdots,\xi_n$ 为独立同分布的正态 $\mathcal{N}(m,\sigma^2)$ 随机变量, 并需利用第 30 题 $(n\geqslant 2)$ 和第 31 题 $(n\geqslant 1)$ 的结果.

设 m 未知, σ^2 已知, 且 $\sigma^2=\sigma_0^2$.

(a) 证明, 对于 $\overline{\xi}=\frac{1}{n}\sum_{k=1}^{n}\xi_k(=\frac{1}{n}T_1(\xi))$, 有

$$\mathbf{E}\overline{\xi}=m\quad(\text{无偏性}),\quad \mathbf{D}\overline{\xi}=\frac{\sigma_0^2}{n}.$$

(b) 证明, (在假设 $\sigma^2=\sigma_0^2$ 之下) 样本均值 $\overline{x}$ 是*有效估计*, 亦即最小方差的无偏估计. 为此, 请证明, 关于参数 m 的无偏估计 $T(x)$ 满足*拉奥–克拉默不等式*:

$$\mathbf{D}T(\xi)\geqslant\frac{1}{n\mathbf{E}\left(\dfrac{\partial\ln p_{(m,\sigma_0^2)}(\xi)}{\partial m}\right)}=\frac{1}{n/\sigma_0^2},$$

其中

$$p_{(m,\sigma_0^2)}(x)=\frac{1}{\sqrt{2\pi\sigma_0^2}}e^{-\frac{(x-m)^2}{2\sigma_0^2}}.$$

(c) 证明, 随机变量

$$\frac{\overline{\xi}-m}{\sigma_0/\sqrt{n}}$$

服从标准正态 $\mathcal{N}(0,1)$ 分布, 且对于 $0<\varepsilon<1$, 当 $\lambda(\varepsilon)$ 满足条件

$$1-\varepsilon=\frac{1}{\sqrt{2\pi}}\int_{-\lambda(\varepsilon)}^{\lambda(\varepsilon)}e^{-t^2/2}dt\quad(=2\Phi(\lambda(\varepsilon))-1)$$

时, 区间

$$\left(\overline{x}-\frac{\sigma_0}{\sqrt{n}}\lambda(\varepsilon),\ \overline{x}+\frac{\sigma_0}{\sqrt{n}}\lambda(\varepsilon)\right)$$

是参数 m 的可靠性系数为 $1-\varepsilon$ 的置信区间, 亦即 "覆盖概率"

$$P_{(m,\sigma_0^2)}\left\{\overline{\xi}-\frac{\sigma_0}{\sqrt{n}}\lambda(\varepsilon)\leqslant m\leqslant\overline{\xi}+\frac{\sigma_0}{\sqrt{n}}\lambda(\varepsilon)\right\}=1-\varepsilon,$$

其中 $P_{(m,\sigma_0^2)}$ 为具有密度函数 $p_{(m,\sigma_0^2)}$ 的分布 (试比较第一章第 7 节第 2 小节中的定义).

33. (关于正态分布 $\mathcal{N}(m,\sigma^2)$ 的统计量. IV: $m=m_0$, σ^2 未知.) 如果 m 已知 ($m=m_0$), 那么就不必以 $s^2(x)=\frac{1}{n}\sum_{k=1}^n(x_k-\overline{x})^2$ 作为 σ^2 的估计, 而只需用

$$s_0^2(x)=\frac{1}{n}\sum_{k=1}^n(x_k-m_0)^2$$

来估计 σ^2 了.

(a) 证明,

$$\mathbf{E}s_0^2(\xi)=\sigma^2\quad(\text{无偏性}),\quad \mathbf{D}s_0^2(\xi)=\frac{2\sigma^4}{n}.$$

(b) 证明, $s_0^2(x)$(在 $m=m_0$ 的情形下) 是 σ^2 的有效估计, 亦即最小方差的无偏估计. 为此, 需要证明, 关于 σ^2 的无偏估计 $T(x)$ 满足如下形式的拉奥–克拉默不等式:

$$\mathbf{D}T(\xi)\geqslant\frac{1}{n\mathbf{E}\left(\dfrac{\partial\ln p_{(m,\sigma_0^2)}(\xi)}{\partial\sigma^2}\right)}=\frac{1}{n/2\sigma^4}.$$

为了解决关于估计量 $s_0^2(x)$ 的精确度问题, 我们利用下述方法构造关于 σ^2 的置信区间:

对 $x=(x_1,\cdots,x_n)$, 令

$$\chi_n^2(x)=\sum_{k=1}^n\left(\frac{x_k-m_0}{\sigma}\right)^2.$$

由于

$$\chi_n^2(\xi)\overset{\mathrm{d}}{=}\sum_{k=1}^n\eta_k^2\quad(=\chi_n^2),$$

故由第 8 节中的 (34) 式, 知 $\chi_n^2(\xi)$ 服从自由度为 n 的 χ^2 分布, 其密度函数为:

$$f_{\chi_n^2}(x)=\frac{x^{n/2-1}e^{-x/2}}{2^{n/2}\Gamma(n/2)},\quad x\geqslant 0$$

(亦可参阅第 3 节表 3).

又由于

$$s_0^2(x)=\frac{\chi_n^2(x)\sigma^2}{n},$$

故知

$$\mathbf{P}_{(m_0,\sigma_0^2)}\left\{\frac{s_0^2(\xi)n}{\sigma^2}\leqslant x\right\}=\int_0^x f_{\chi_n^2}(t)dt.$$

(关于 χ^2 分布, 有关于 $n=1,2,\cdots$ 的许多分布值表格.) 所以, 对于给定的 $0<\varepsilon<1$, 可以找到 $\lambda'(\varepsilon)$ 和 $\lambda''(\varepsilon)$, 使得

$$\int_0^{\lambda'(\varepsilon)}f_{\chi_n^2}(t)dt=\frac{\varepsilon}{2},\quad\int_{\lambda''(\varepsilon)}^\infty f_{\chi_n^2}(t)dt=\frac{\varepsilon}{2}.$$

此时就有

$$\int_{\lambda'(\varepsilon)}^{\lambda''(\varepsilon)}f_{\chi_n^2}(t)dt=1-\varepsilon,$$

并且由于

$$\left\{\frac{s_0^2(x)n}{\lambda''(\varepsilon)}\leqslant\sigma^2\leqslant\frac{s_0^2(x)n}{\lambda'(\varepsilon)}\right\}=\{\lambda'(\varepsilon)\leqslant\chi_n^2(x)\leqslant\lambda''(\varepsilon)\},$$

所以区间

$$\left(\frac{s_0^2(x)n}{\lambda''(\varepsilon)},\ \frac{s_0^2(x)n}{\lambda'(\varepsilon)}\right)$$

就是关于 σ^2 的置信系数为 $1-\varepsilon$ 的置信区间.

假设 $a'(\varepsilon)$ 与 $a''(\varepsilon)$ 被选取得使有

$$\int_{a'(\varepsilon)}^{a''(\varepsilon)}f_{\chi_n^2}(t)dt=1-\varepsilon.$$

易知, 这种根据给定的 $\varepsilon>0$, 所选取的 $a'(\varepsilon)$ 与 $a''(\varepsilon)$ 并不是唯一的.

(c) 试问, 对于怎样的 $a'(\varepsilon)$ 与 $a''(\varepsilon)$, 可以得到关于 σ^2 的最窄的置信系数为 $1-\varepsilon$ 的置信区间? 它们是否重合于按照上面的办法所选取的 $\lambda'(\varepsilon)$ 与 $\lambda''(\varepsilon)$?

34. (关于正态分布 $\mathcal{N}(m,\sigma^2)$ 的统计量. V: m 未知, σ^2 未知.)

(a) 证明, 在所考察的情形下, 对于 $n>1$,

$$\overline{x}=\frac{1}{n}\sum_{k=1}^n x_k,\quad s_1^2(x)\equiv\frac{n}{n-1}s^2(x)=\frac{1}{n-1}\sum_{k=1}^n(x_k-\overline{x})^2$$

分别是关于 m 和 σ^2 的无偏估计.

(b) 证明, 统计量

$$t_{n-1}(x)=\frac{\overline{x}-m}{s_1(x)/\sqrt{n}}$$

具有自由度为 $n-1$ 的学生分布.

提示: 将 $t_{n-1}(x)$ 表示为:

$$t_{n-1}(x)=\frac{\dfrac{\overline{x}-m}{\sigma}\sqrt{n}}{s_1(x)/\sigma},$$

然后分别指出:

(i) 分子服从标准正态分布 $\mathcal{N}(0,1)$;

(ii) 分母 $\frac{s_1(\xi)}{\sigma}$ 的分布与 $\sqrt{\frac{1}{n-1}\chi^2_{n-1}}$ 相同, 其中 $\chi^2_{n-1}\overset{\mathrm{d}}{=}\sum_{k=1}^{n-1}\eta_k$, 而 $\eta_1,\cdots,\eta_{n-1}$ 是相互独立的标准正态 $\mathcal{N}(0,1)$ 随机变量;

(iii) 随机变量 $\frac{\bar{\xi}-m}{\sigma}\sqrt{n}$ 与 $\frac{s_1(\xi)}{\sigma}$ 相互独立.

由 (i)(ii)(iii) 和第 8 节中的 (38) 式, 即可推出关于随机变量 $t_{n-1}(\xi)$ 的分布.

(c) 试利用 $t_{n-1}(\xi)$ 的分布, 以及 $t_{n-1}(x)=\frac{\overline{x}-m}{s_1(x)/\sqrt{n}}$ 的表达式, 构造关于 m 的置信系数为 $1-\varepsilon$ 的置信区间.

(d) 首先证明随机变量 $(n-1)\left(\frac{s_1(\xi)}{\sqrt{n}}\right)^2$ 服从自由度为 $n-1$ 的 χ^2 分布, 再利用这一结果构造关于 σ 的置信系数为 $1-\varepsilon$ 的置信区间.

35. 设 $\varphi(t)$ 为第 28 题中的特征函数. 证明, 如果 $0<a_1<\cdots<a_n$, 而 $p_k\geqslant 0$, 使得 $\sum\limits_{k=1}^{n}p_k=1$, 则函数

$$\psi(t)=\sum_{k=1}^{n}p_k\varphi\left(\frac{t}{a_k}\right)$$

亦为特征函数.

36. 如果高斯序列 $X=(X_n)_{n\geqslant 0}$ 分别具有如下形式的协方差函数:

$$e^{-|i-j|}\quad 和\quad \min(i,j)\ \ (=2^{-1}(|i|+|j|-|i-j|)),\quad i,j=0,1,2,\cdots,$$

那么它具有怎样的结构性质 (独立增量, 平稳, 马尔可夫等等)?

37. 设 N 为标准正态随机变量 ($N\sim\mathcal{N}(0,1)$). 证明, 对于 $\alpha<1$, 有

$$\mathbf{E}\frac{1}{|N|^\alpha}=\frac{1}{\sqrt{2^\alpha\pi}}\Gamma\left(\frac{1}{2}-\frac{\alpha}{2}\right).$$

38. 设 X 与 Y 为相互独立的标准正态 $\mathcal{N}(0,1)$ 随机变量. 证明,

$$\mathbf{E}\left(\frac{1}{(X^2+Y^2)^{p/2}}\right)<\infty,$$

当且仅当 $p<2$.

39. 设第 38 题中的条件成立, 而

$$T=\frac{1}{2}(X^2+Y^2),\quad g=\frac{X^2}{X^2+Y^2}.$$

证明,

(a) T 与 g 相互独立.

(b) T 服从指数分布 ($\mathbf{P}\{T>t\}=e^{-t},\ t>0$).

(c) g 服从反正弦分布(密度函数为: $\frac{1}{\pi\sqrt{x(1-x)}},\ x\in(0,1)$).

40. 设 $B=(B_t)_{t\geqslant 0}$ 为布朗运动, 而

$$T_a=\min\{t\geqslant 0:\ B_t=a\}$$

为其首次达到水平 $a>0$ 的时刻. (如果 $\{\cdot\}=\varnothing$, 则令 $T_a=\infty$.)

试利用关于布朗运动的反射原理($\mathbf{P}\left\{\sup\limits_{s\leqslant t}B_s>a\right\}=2\mathbf{P}\{B_t\geqslant a\}$, 参阅例如, 参考文献 [17], 第三章第 10 节, p89-90), 证明, 密度函数 $p_a(t)=\frac{\partial\mathbf{P}\{T_a\leqslant t\}}{\partial t}$, $t>0$, 为

$$p_a(t)=\frac{a}{\sqrt{2\pi t^3}}e^{-a^2/(2t)}.$$

提示: 利用关系式 $\mathbf{P}\{T_a\leqslant t\}=2\mathbf{P}\{B_t\geqslant a\}$.

41. 设 $T=T_1$, 其中 T_a 如上面的第 40 题所定义. 证明

$$T\overset{\text{law}}{=}N^{-2},$$

其中 N 为标准高斯随机变量 ($N\sim\mathcal{N}(0,1)$). 再证明, T 的拉普拉斯变换和傅立叶变换分别为:

$$\mathbf{E}e^{-\frac{\lambda^2}{2}T}=\mathbf{E}e^{-\frac{\lambda^2}{2}N^{-2}}=e^{\lambda},\quad \lambda\geqslant 0,$$

$$\mathbf{E}e^{itT}=\mathbf{E}e^{it\frac{1}{N^2}}=\exp\left\{-|t|^{1/2}\left(1-i\frac{t}{|t|}\right)\right\},\quad t\in\mathbb{R}.$$

事实上, 随机变量 $1/N^2$ 从结构上可以视为具有参数为 $\alpha=1/2$, $\beta=0$, $\theta=-1$, $d=1$ 的稳定分布 (参阅第三章第 6 节中的 (9) 与 (10) 式).

42. 设 X 与 Y 为相互独立的正态 $\mathcal{N}(0,\sigma^2)$ 随机变量.

(a) 证明

$$\frac{2XY}{\sqrt{X^2+Y^2}}\quad 与\quad \frac{X^2-Y^2}{\sqrt{X^2+Y^2}}$$

为相互独立的均值为 0, 方差为 1/2 的正态分布随机变量.

(b) 由此推出

$$\frac{X}{Y}-\frac{Y}{X}\overset{\text{law}}{=}2C,$$

其中 C 为具有密度函数 $1/(\pi(1+x^2))$ 的柯西分布随机变量 ($x\in\mathbb{R}$), 并且

$$C-\frac{1}{C}\overset{\text{law}}{=}2C.$$

(c) 推广上述性质: 对任何 $a>0$, 都有

$$C-\frac{a}{C}\overset{\text{law}}{=}(1+a)C.$$

(d) 证明, 随机变量 X^2+Y^2 与 $\frac{X}{\sqrt{X^2+Y^2}}$ 相互独立.

提示: (a) 利用第 8 节第 13 题中关于 X 和 Y 的表达式.

(b) 利用第 8 节中第 27 题 (a) 中的结论.

(c) 为证明所需结论, 只需证明, 如果 $f(x)=\frac{x-ax^{-1}}{1+a}$, 则对任何有界函数 $g(x)$, 都有积分 $\int_{-\infty}^{\infty}g(f(x))\frac{dx}{1+x^2}$ 与积分 $\int_{-\infty}^{\infty}g(x)\frac{dx}{1+x^2}$ 相等.

43. 证明, 函数

$$R(s,t)=\frac{1}{2}(t^{2H}+s^{2H}-|t-s|^{2H})$$

为非负定的 (参阅 (24) 式), 其中 $0<H\leqslant 1$, $s,t\geqslant 0$. 并且存在 (零均值的) 高斯过程 $B^H=(B_t^H)_{t\geqslant 0}$ 以 $R(s,t)$ 作为自己的协方差函数. (由周知的关于连续修正存在性的柯尔莫戈洛夫定理, 例如, 参阅 [17], 可以断言, 能够将这样的过程 $B^H=(B_t^H)_{t\geqslant 0}$ 取为轨道连续的过程. 该过程通常称为具有哈尔斯特参数 H 的断口 (分数) 布朗运动.

再证明, 所述的函数 $R(s,t)$ 在 $H>1$ 时不是非负定的.

44. (达尔穆亚–斯基托维奇定理) 设 $\xi_1,\cdots,\xi_n$ 为独立同分布的随机变量, 而 $a_1,\cdots,a_n,b_1,\cdots,b_n$ 为非零常数. 证明如下的刻画性定理: 如果随机变量 $\sum\limits_{i=1}^n a_i\xi_i$ 与 $\sum\limits_{i=1}^n b_i\xi_i$ 相互独立, 则随机变量 $\xi_1,\cdots,\xi_n$ 有正态分布. (在 $n=2$, $a_1=a_2=1$, $b_1=1$, $b_2=-1$ 时, 即为第 10 题中的伯恩斯坦定理.)

45. 设 $\xi_1,\xi_2,\cdots$ 为相互独立的标准正态 $\mathscr{N}(0,1)$ 随机变量序列. 证明, 随机变量

$$X_n=\sqrt{n}\frac{\sum\limits_{i=1}^n\xi_i}{\sum\limits_{i=1}^n\xi_i^2}\quad 与\quad Y_n=\frac{\sum\limits_{i=1}^n\xi_i}{\left(\sum\limits_{i=1}^n\xi_i^2\right)^{1/2}}$$

(在 $n\to\infty$ 时) 依分布收敛到正态随机变量 $\mathscr{N}(0,1)$.

46. 设 (X,Y) 是具有 $\mathbf{E}X=\mathbf{E}Y=0$, $\mathbf{D}X=\mathbf{D}Y=1$ 的高斯随机变量对, 相关系数为 $\rho_{X,Y}$. 设 $\Phi(x)$ 为标准正态分布函数, 即 $\Phi(x)=(2\pi)^{-1/2}\int_{-\infty}^{x}e^{-y^2/2}dy$. 证明, 随机变量 $\Phi(X)$ 与 $\Phi(Y)$ 的相关系数 $\rho_{\Phi(X),\Phi(Y)}$ 等于

$$\rho_{\Phi(X),\Phi(Y)}=\frac{6}{\pi}\arcsin\frac{\rho_{X,Y}}{2}.$$

47. 设 (X,Y,Z) 为三维高斯随机变量, 有 $\mathbf{E}X=\mathbf{E}Y=\mathbf{E}Z=0$, $\mathbf{D}X=\mathbf{D}Y=\mathbf{D}Z=1$ 以及相关系数 $\rho(X,Y)=\rho_1$, $\rho(X,Z)=\rho_2$, $\rho(Y,Z)=\rho_3$. 证明 (试比较第 23 题中的断言 (b))

$$\mathbf{P}\{X>0,\ Y>0,\ Z>0\}=\frac{1}{8}+\frac{1}{4\pi}(\arcsin\rho_1+\arcsin\rho_2+\arcsin\rho_3).$$

提示: 令 $A=\{X>0\}$, $B=\{Y>0\}$, $C=\{Z>0\}$, 记 $p=\mathbf{P}(A\cap B\cap C)$, 则根据容斥公式 (第一章第 1 节第 12 题), 得

$$\begin{aligned}1-p&=\mathbf{P}(A\cup B\cup C)\\&=[\mathbf{P}(A)+\mathbf{P}(B)+\mathbf{P}(C)]-[\mathbf{P}(A\cap B)+\mathbf{P}(A\cap C)+\mathbf{P}(B\cap C)]+p.\end{aligned}$$

再利用断言 23(b).

48. 设 $B^\circ=(B_t^\circ)_{0\leqslant t\leqslant 1}$ 为布朗桥, 以 $\mathscr{R}$ 表示其 "振幅", 即

$$\mathscr{R}=\sqrt{\frac{2}{\pi}}\left(\max_{0\leqslant t\leqslant 1}B_t^\circ-\min_{0\leqslant t\leqslant 1}B_t^\circ\right).$$

证明, B° 振幅的平方 $\mathscr{R}^2$ 的拉普拉斯变换 $\mathbf{E}e^{-\lambda\mathscr{R}^2}$, $\lambda>0$, 为

$$\mathbf{E}e^{-\lambda\mathscr{R}^2}=\left(\frac{\sqrt{\lambda\pi}}{\operatorname{sh}\sqrt{\lambda\pi}}\right)^2.$$

49. (维斯可夫) 设 η 与 ζ 为相互独立的标准正态 $\mathscr{N}(0,1)$ 随机变量. 证明,

(a) 对于任何满足条件 $\mathbf{E}|f(x+(\eta+i\zeta))|<\infty$ 的函数 $f=f(z)$, $z\in\mathbb{C}$, 都有如下的 "求均值" 性质:

$$f(x)=\mathbf{E}f(x+(\eta+i\zeta)).$$

(b) 对于埃尔米特多项式 $\mathrm{He}_n(x)$, $n\geqslant 0$ (参阅附录 p327), 都有 $\mathrm{He}_n(x)=\mathbf{E}(x+i\zeta)^n$ 和 $\mathbf{E}\,\mathrm{He}_n(x+\eta)=x^n$.

第三章　概率测度的接近程度和收敛性. 中心极限定理

§1. 概率测度和分布的弱收敛

1. 称定义在 $\mathbb{R}^m$ 上的函数 $F = F(x)$ 在 $x \in \mathbb{R}^m$ 处连续, 如果对任给的 $\varepsilon > 0$, 都存在 $\delta > 0$, 使得 $|F(x) - F(y)| < \varepsilon$ 对一切满足 $x - \delta e < y < x + \delta e$ 的 $y \in \mathbb{R}^m$ 成立, 其中 $e = (1, \cdots, 1) \in \mathbb{R}^m$. 称分布函数序列 $(F_n)_{n \geqslant 1}$ 淡收敛于 F(记为 $F_n \Rightarrow F$) ①, 如果当 $n \to \infty$ 时, 在 $F = F(x)$ 的任一连续点 $x \in \mathbb{R}^m$ 处都有 $F_n(x) \to F(x)$.

证明, 定理 2 的断言在空间 $\mathbb{R}^m (m > 1)$ 中仍然成立. (参阅定理 1 的注 1.)

提示: 只需证明 (1) $\Leftrightarrow$ (4). 为证 (1) $\Rightarrow$ (4), 设 $x \in \mathbb{R}^m$ 是 F 的连续点. 证明, 如果 $\partial(-\infty, x]$ 是集合 $(-\infty, x] = (-\infty, x_1] \times \cdots \times (-\infty, x_m]$ 的边界, 则有 $\mathbf{P}(\partial(-\infty, x]) = 0$, 因而就有 $\mathbf{P}_n((-\infty, x]) \to \mathbf{P}((-\infty, x])$, 亦即 $F_n(x) \to F(x)$. 为证 m 中的反向蕴涵关系 (4) $\Rightarrow$ (1), 只需仿照 1 维情形下定理 2 的有关证明.

2. 证明, 在空间 $\mathbb{R}^m$ 中, “基本集合” 类 $\mathscr{K}$ 就是收敛确定类.

3. 设 E 是空间 $\mathbb{R}^\infty$, C 或 D 之一 (参阅第二章第 2 节). 称 (定义于由全体开集所生成的 σ-代数 $\mathscr{E}$ 上的) 概率测度序列 $(\mathbf{P}_n)_{n \geqslant 1}$ 在有限维分布的意义上淡收敛到测度 $\mathbf{P}$ (记为 $\mathbf{P}_n \overset{f}{\Rightarrow} \mathbf{P}$), 如果当 $n \to \infty$ 时, 对一切满足条件 $\mathbf{P}(\partial A) = 0$ 的柱集A, 都有 $\mathbf{P}_n(A) \to \mathbf{P}(A)$.

①原文为: Сходится в основном, 直译应为基本收敛, 此处根据国内的通行称呼, 译为淡收敛 —— 译者注.

证明, 对于空间 $\mathbb{R}^\infty$, 有

$$(\mathbf{P}_n \overset{f}{\Rightarrow} \mathbf{P}) \quad \Leftrightarrow \quad (\mathbf{P}_n \Rightarrow \mathbf{P}). \tag{$*$}$$

试问, 该结论对于空间 C 与 D 是否也成立?

提示: $(*)$ 式中的蕴涵关系 $\Leftarrow$ 是显然的, 我们只需证明 $(\mathbf{P}_n \overset{f}{\Rightarrow} \mathbf{P}) \Rightarrow (\mathbf{P}_n \Rightarrow \mathbf{P})$. 设 f 是 $C(\mathbb{R}^\infty)$ 中的有界函数 ($|f| \leqslant c$). 对于 $m \in N = \{1, 2, \cdots\}$, 我们定义 $f_m: \mathbb{R}^\infty \to \mathbb{R}$ 如下:

$$f_m(x_1, \cdots, x_m, x_{m+1}, \cdots) = f(x_1, \cdots, x_m, 0, 0, \cdots).$$

则有 $f_m \in C(\mathbb{R}^\infty)$, $|f_m| \leqslant c$, 并且对任何 $x \in \mathbb{R}^\infty$, 都有 $f_m(x) \to f(x)$. 考察集合

$$A_m = \{x \in \mathbb{R}^\infty : \ |f_m(x) - f(x)| \leqslant \varepsilon\},$$

证明, 对于充分大的 n 和 m, 都有

$$\left| \int_{\mathbb{R}^\infty} f_m d\mathbf{P}_n - \int_{\mathbb{R}^\infty} f d\mathbf{P}_n \right| \leqslant \varepsilon \mathbf{P}_n(A_m) + 2c\mathbf{P}_n(\overline{A}_m) \leqslant \varepsilon + 4c\varepsilon.$$

再利用对于每个 m 都有 $\int_{\mathbb{R}^\infty} f_m d\mathbf{P} \to \int_{\mathbb{R}^\infty} f d\mathbf{P}$ 的事实, 即可由上述事实推出, 对于充分大的 n, 有

$$\left| \overline{\lim_n} \int_{\mathbb{R}^\infty} f d\mathbf{P}_n - \int_{\mathbb{R}^\infty} f d\mathbf{P} \right| \leqslant \varepsilon + 4c\varepsilon,$$

$$\left| \underline{\lim_n} \int_{\mathbb{R}^\infty} f d\mathbf{P}_n - \int_{\mathbb{R}^\infty} f d\mathbf{P} \right| \leqslant \varepsilon + 4c\varepsilon.$$

由勒贝格控制收敛定理, 知 $\int_{\mathbb{R}^\infty} f_m d\mathbf{P} \to \int_{\mathbb{R}^\infty} f d\mathbf{P}$, 再由上面两个不等式, 即得

$$\left| \overline{\lim_n} \int_{\mathbb{R}^\infty} f d\mathbf{P}_n - \int_{\mathbb{R}^\infty} f d\mathbf{P} \right| \leqslant \varepsilon + 4c\varepsilon,$$

$$\left| \underline{\lim_n} \int_{\mathbb{R}^\infty} f d\mathbf{P}_n - \int_{\mathbb{R}^\infty} f d\mathbf{P} \right| \leqslant \varepsilon + 4c\varepsilon.$$

由 $\varepsilon > 0$ 的任意性, 即可断言

$$\int_{\mathbb{R}^\infty} f d\mathbf{P}_n \to \int_{\mathbb{R}^\infty} f d\mathbf{P}, \quad n \to \infty.$$

4. 设 F 与 G 为数轴上的分布函数, 而

$$L(F, G) = \inf \{h > 0 : \ F(x - h) - h \leqslant G(x) \leqslant F(x + h) + h\}$$

是 (F 与 G 之间的) 莱维距离. 证明, 淡收敛与依莱维距离收敛相互等价, 亦即

$$(F_n \Rightarrow F) \quad \Leftrightarrow \quad (L(F_n, F) \to 0).$$

提示: 蕴涵关系 $(L(F_n, F) \to 0) \Rightarrow (F_n \Rightarrow F)$ 可以立即由定义推出. 反过来的蕴涵关系要用反证法证明, 假设 $F_n \Rightarrow F$, 但是却有 $L(F_n, F) \not\to 0$, 然后找出矛盾来.

5. 设 $F_n \Rightarrow F$, 而分布函数 F 连续. 证明, 此时 $F_n(x)$ 向 $F(x)$ 的收敛对一切 $x \in \mathbb{R}$ 一致成立, 亦即

$$\sup_x |F_n(x) - F(x)| \to 0, \quad n \to \infty.$$

提示: 给定 $\varepsilon > 0$, 取 $m > 1/\varepsilon$. 根据 F 的连续性, 可以找到 $x_1, \cdots, x_{m-1}$, 使得 $F(x_i) = i/m$; 并且对于充分大的 n, 有 $|F_n(x_i) - F(x_i)| < \varepsilon$, $i = 1, \cdots, m-1$. 这样一来, 就对任何 $x \in [x_k, x_{k+1}]$(其中 $x_0 = -\infty$, $x_m = \infty$), 都有

$$F_n(x) - F(x) \leqslant F_n(x_{k+1}) - F(x_k) \leqslant F(x_{k+1}) + \varepsilon - F(x_k) = \varepsilon + \frac{1}{m} < 2\varepsilon.$$

同理有 $F(x) - F_n(x) < 2\varepsilon$. 此即表明, 对任何 $x \in \mathbb{R}$, 都有 $|F(x) - F_n(x)| < 2\varepsilon$.

6. 证明, 关于定理 1 的注 1 中的断言.

7. 证明, 关于定理 1 的注 2 中的条件 (I*) ~ (IV*) 的相互等价性.

8. 证明, $\mathbf{P}_n \xrightarrow{w} \mathbf{P}$ 当且仅当, 序列 $(\mathbf{P}_n)$ 的任一子列 $(\mathbf{P}_{n'})$ 中, 都包含着子列 $(\mathbf{P}_{n''})$, 使得 $\mathbf{P}_{n''} \xrightarrow{w} \mathbf{P}$.

提示: 必要性显然. 为证充分性, 需要指出, 假若 $\mathbf{P}_n \stackrel{w}{\not\to} \mathbf{P}$, 则可找到有界函数 f, $\varepsilon > 0$ 和子列 (n'), 使得

$$\left|\int_E f d\mathbf{P}_{n'} - \int_E f d\mathbf{P}\right| > \varepsilon.$$

此与存在 $(n'') \subseteq (n')$, 使得 $\mathbf{P}_{n''} \xrightarrow{w} \mathbf{P}$ 的事实相矛盾.

9. 试举例说明, 存在 $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ 上的概率测度 P, P_n, $n \geqslant 1$, 对其有 $P_n \xrightarrow{w} P$, 但是对任何 $B \in \mathscr{B}(\mathbb{R})$, 却都有 $P_n(B) \to P(B)$ 不成立.

10. 试举例说明, 存在分布函数 $F = F(x)$, $F_n = F_n(x)$, $n \geqslant 1$, 对其有 $F_n \xrightarrow{w} F$, 但是却有 $\sup_x |F_n(x) - F(x)| \not\to 0$, $\quad n \to \infty$.

11. 在关于概率论的许多指导书籍中, 有关分布函数 $F_n\ (n \geqslant 1)$ 收敛的定理 2 中的命题 (4) ⇒ (3) 往往与赫利以及布雷 (Bray) 的名字联系在一起. 这种联系导致对如下命题的证明,

(a) 赫利–布雷引理: 如果 $F_n \Rightarrow F$(参阅定义 1), 则

$$\lim_n \int_a^b g(x) dF_n(x) = \int_a^b g(x) dF(x),$$

其中 a 与 b 均为分布函数 $F = F(x)$ 的连续点, 而 $g = g(x)$ 是区间 $[a, b]$ 上的连续函数.

(b) 赫利–布雷定理: 如果 $F_n \Rightarrow F$, 而 $g = g(x)$ 是 $\mathbb{R}$ 中的连续函数, 则

$$\lim_n \int_{-\infty}^{\infty} g(x) dF_n(x) = \int_{-\infty}^{\infty} g(x) dF(x).$$

12. 设 $F_n \Rightarrow F$, 对某个 $b > 0$, 序列 $\left(\int |x|^b dF_n(x)\right)_{n \geqslant 1}$ 有界. 证明, 此时有 (其中 $\left(\int (\cdot) = \int_{-\infty}^{\infty} (\cdot)\right)$):

$$\lim_n \int |x|^a dF_n(x) = \int |x|^a dF(x), \quad 0 \leqslant a \leqslant b,$$

$$\lim_n \int x^k dF_n(x) = \int x^k dF(x), \quad \text{对一切 } k = 1, 2, \cdots, [b],\ k \neq b.$$

13. 设 $F_n \Rightarrow F$, 而 $\mu = \operatorname{med}(F)$ 与 $\mu_n = \operatorname{med}(F_n)$ 分别为 F 和 F_n 的中位数 (参阅第一章第 4 节第 5 题). 假设中位数 μ 和 μ_n 对于所有 $n \geqslant 1$ 均唯一确定. 证明, $\mu_n \to \mu$.

14. 设分布函数 F 被自己的矩序列 $m^{(k)} = \int_{-\infty}^{\infty} x^k dF(x)$, $k = 1, 2, \cdots$, 唯一确定. 假设 $(F_n)_{n \geqslant 1}$ 为分布函数序列, 使得

$$m_n^{(k)} = \int_{-\infty}^{\infty} x^k dF_n(x) \to m^{(k)} = \int_{-\infty}^{\infty} x^k dF(x), \quad k = 1, 2, \cdots.$$

证明, $F_n \Rightarrow F$.

15. 设 μ 是距离空间 $(E, \mathscr{E}, \rho)$ 中的 σ-有限测度, 其中 $\mathscr{E}$ 是由博雷尔集合构成的 σ-代数. 证明, 对一切 $B \in \mathscr{E}$, 都有

$$\mu(B) = \sup_{F \subseteq B} \mu(F) = \inf_{G \supseteq B} \mu(G),$$

其中, F 与 G 分别为 $\mathscr{E}$ 中的闭集与开集.

16. 证明, 数轴 $\mathbb{R}$ 上的分布函数序列 $F_n\ (n \geqslant 1)$ 弱收敛到 F ($F_n \xrightarrow{w} F$), 当且仅当, 在 $\mathbb{R}$ 存在一个处处稠密的集合 D, 使得对任何 $x \in D$, 都有 $F_n(x) \to F(x)$.

17. 设 $g(x)$ 与 $(g_n(x))_{n \geqslant 1}$, $x \in \mathbb{R}$, 为连续函数, 使得

$$\sup_{x, n} |g_n(x)| \leqslant c < \infty,$$

且对任何有界区间 $B = [a, b]$, 都有

$$\lim_n \sup_{x \in B} |g_n(x) - g(x)| = 0.$$

证明, 如果成立分布函数的收敛关系 $F_n \Rightarrow F$, 则有

$$\lim_n \int_{\mathbb{R}} g_n(x) dF_n(x) = \int_{\mathbb{R}} g(x) dF(x).$$

并举例说明, 对于该收敛性, 仅仅有逐点收敛 $g_n(x) \to g(x)$, $x \in \mathbb{R}$, 通常是不够的.

18. 设分布函数序列 $F_n \Rightarrow F$, $n \to \infty$.

(a) 试举例说明, 一般来说,

$$\int_{\mathbb{R}} x dF_n(x) \not\to \int_{\mathbb{R}} x dF(x).$$

(b) 证明, 如果对某个 $k \geqslant 1$, 有 $\sup\limits_n \int_{\mathbb{R}} |x|^k dF_n(x) \leqslant c < \infty$, 则对一切 $1 \leqslant l \leqslant k-1$, 都有

$$\int_{\mathbb{R}} x^l dF_n(x) \to \int_{\mathbb{R}} x^l dF(x).$$

19. 推广上题, 证明, 如果 $f = f(x)$ 为连续函数, 未必有界, 但具有如下性质:

$$\lim_{|x|\to\infty} \frac{|f(x)|}{g(x)} = 0,$$

其中 $g = g(x)$ 是正值函数, 使得 $\sup\limits_n \int_{\mathbb{R}} g(x) dF_n(x) \leqslant c < \infty$, 则有

$$\int_{\mathbb{R}} f(x) dF_n(x) \to \int_{\mathbb{R}} f(x) dF(x).$$

§2. 概率分布族的相对紧性和稠密性

1. 对空间 $\mathbb{R}^n$, $n \geqslant 2$, 证明相应的定理 1 和定理 2.

2. 设 $\mathbf{P}_\alpha$ 是数轴上的参数为 m_α 与 $\sigma_\alpha^2 (\alpha \in \mathfrak{u})$ 的高斯测度. 证明, 族 $\mathscr{P} = \{\mathbf{P}_\alpha : \alpha \in \mathfrak{u}\}$ 是稠密的, 当且仅当, 存在 a 与 b, 使得

$$|m_\alpha| \leqslant a, \quad \sigma_\alpha^2 \leqslant b, \quad \alpha \in \mathfrak{u}.$$

提示: 充分性的证明: 对任何 $\alpha \in \mathfrak{u}$, 均可找到随机变量 $\eta_\alpha \sim \mathcal{N}(0,1)$, 使得 $\xi_\alpha = m_\alpha + \sigma_\alpha \eta_\alpha$. 将此与题中条件相结合, 得知 $\mathbf{P}\{|\xi_\alpha| \geqslant n\} \leqslant \mathbf{P}\left\{|\eta_\alpha| \geqslant \frac{n-a}{b}\right\}$, 由此即得 $\{\mathbf{P}_\alpha\}$ 的稠密性. 必要性可用反证法证明.

3. 试分别举出定义在 $(\mathbb{R}^\infty, \mathscr{B}(\mathbb{R}^\infty))$ 上的稠密的和非稠密的概率测度族 $\mathscr{P} = \{\mathbf{P}_\alpha : \alpha \in \mathfrak{u}\}$ 的例子.

提示: 考察下列概率测度族:

(a) $\{\mathbf{P}_\alpha\}$, 其中 $\mathbf{P}_\alpha \equiv \mathbf{P}$, 而

$$\mathbf{P}(A) = \begin{cases} 1, & \text{如果 } (0,0,0,\cdots) \in A, \\ 0, & \text{如果 } (0,0,0,\cdots) \notin A. \end{cases}$$

(b) $\{\mathbf{P}_n : n \in \mathbb{N}\}$, 其中 $\mathbf{P}_n$ 为集中在点 $x_n = (n,0,0,\cdots)$ 上的概率测度.

4. 设 P 是距离空间 $(E, \mathscr{E}, \rho)$ 中的概率测度. 称 P 是稠密的 [①], 如果对任何 $\varepsilon > 0$, 都存在紧集 $K \subseteq E$, 使得 $P(K) \geqslant 1-\varepsilon$. 证明如下结果 ("乌拉姆定理"): 波利希空间 (即完备可分的距离空间) 中的任何概率测度 P 都是稠密的.

5. 设 $X = \{X_\alpha;\ \alpha \in \mathfrak{u}\}$ 是某个随机向量族 $(X_\alpha \in \mathbb{R}^d,\ \alpha \in \mathfrak{u})$, 对某 $r > 0$, 有 $\sup\limits_\alpha \mathbf{E}\|X_\alpha\|^r < \infty$. 证明, 分布 $P_\alpha = \mathrm{Law}(X_\alpha)$ 的族 $\mathscr{P} = \{P_\alpha;\ \alpha \in \mathfrak{u}\}$ 是稠密的.

6. 证明, 对于随机向量族 $\{\xi_t,\ t \in T\}$, 其中随机向量 ξ_t 取值于空间 $\mathbb{R}^n$, 稠密性条件等价于条件

$$\lim_{a\to\infty} \sup_{t\in T} \mathbf{P}\{\|\xi_t\| > a\} = 0.$$

(a) 证明, 对于随机向量族 $\{\xi_k,\ k \geqslant 0\}$, 其中随机向量 ξ_k 取值于空间 $\mathbb{R}^n$, 稠密性条件等价于条件

$$\lim_{a\to\infty} \varlimsup_{k\to\infty} \mathbf{P}\{\|\xi_k\| > a\} = 0.$$

(b) 证明, 如果 $\{\xi_k,\ k \geqslant 0\}$ 为非负随机变量序列, 则稠密性条件等价于条件

$$\lim_{\lambda\downarrow 0} \varlimsup_{n\to\infty} (1 - \mathbf{E}e^{-\lambda\xi_n}) = 0.$$

7. 设 $(\xi_k)_{k\geqslant 0}$ 为取值于 $\mathbb{R}^n$ 的随机向量序列, 且 $\xi_k \xrightarrow{\mathrm{d}} \xi$, 即随机向量 ξ_k 的分布 F_k 弱收敛 (等价于淡收敛) 到随机向量 ξ 的分布 F. 证明, 族 $\{\xi_k,\ k \geqslant 0\}$ 是稠密的.

8. 设 $(\xi_k)_{k\geqslant 0}$ 为稠密的随机变量序列, 而对序列 $(\eta_k)_{k\geqslant 0}$, 有 $\eta_k \xrightarrow{\mathbf{P}} 0$, $k \to \infty$. 证明, $\xi_k\eta_k \xrightarrow{\mathbf{P}} 0$, $k \to \infty$.

9. 设 $X_1, X_2, \cdots$ 为可交换随机变量的无穷序列 (参阅第二章第 5 节第 4 题中的定义), 并设 X_i 仅取 0 和 1 两个值.

证明如下结果的正确性: 存在区间 $[0,1]$ 上的一个分布 $G = G(\lambda)$, 使得对任何 $n \geqslant 1$ 及任何 $0 \leqslant k \leqslant n$, 都有

$$\mathbf{P}\{X_1 = 1, \cdots, X_k = 1, X_{k+1} = 0, \cdots, X_n = 0\} = \int_0^1 \lambda^k (1-\lambda)^{n-k} dG(\lambda).$$

(该结论是著名的费内提定理的特例, 该定理断言: 任何可交换随机变量的无穷序列都是若干个独立同分布的随机变量序列的混合; 参阅 [33] 和 [2].)

提示: 令 $A_k = \{X_1 = 1, \cdots, X_k = 1, X_{k+1} = 0, \cdots, X_n = 0\}$, 将概率 $\mathbf{P}(A_k)$ 表示为

$$\mathbf{P}(A_k) = \sum_{j=0}^{m} \mathbf{P}(A_k \mid S_m = j)\mathbf{P}\{S_m = j\}, \tag{$*$}$$

① 原文为: плотной, 直译为稠密的, 但根据国际上和我国的通行叫法, 应当译为胎紧, 现为保持原书面目, 译为稠密的 —— 译者注.

其中 $m \geqslant n$, 而 $S_m = \sum_{i=1}^{m} X_i$. 再利用可交换性, 证明, 可将 (*) 式的右端写为如下形式:

$$\mathbf{E} \prod_{i=1}^{k-1}(mY_m - i) \cdot \prod_{j=0}^{n-k-1}(m(1-Y_m)-j) \cdot \frac{1}{\prod_{l=0}^{n-1}(m-l)},$$

其中 $Y_m = \frac{S_m}{m}$. (对于充分大的 m, 该式近似于 $\mathbf{E}[Y_m^k(1-Y_m)^{n-k}]$.) 然后, 令 $m \to \infty$, 并证明其极限可表示为 $\int_0^1 \lambda^k(1-\lambda)^{n-k}dG(\lambda)$, 其中 $G(\lambda)$ 是区间 $[0,1]$ 上的某个分布.

10. 设 $\xi_1, \cdots, \xi_n$ 是仅取 0 和 1 两个值的可交换随机变量. 证明,

(a) $\mathbf{P}(\xi_i = 1 \,|\, S_n) = \frac{S_n}{n}$, 其中 $S_n = \xi_1 + \cdots + \xi_n$.

(b) $\mathbf{P}(\xi_i = 1, \xi_j = 1 \,|\, S_n) = \frac{S_n(S_n-1)}{n(n-1)}$, 其中 $i \neq j$.

11. 为推广第一章第 11 节中的定理 2, 证明, 如果 $\eta_1, \cdots, \eta_n$ 是取值于 $\{0,1,2,\cdots\}$ 的可交换随机变量, 而 $S_k = \eta_1 + \cdots + \eta_k$, $1 \leqslant k \leqslant n$, 则

$$\mathbf{P}(S_k < k \text{ 对一切 } 1 \leqslant k \leqslant n \,|\, S_n) = \left(1 - \frac{S_n}{n}\right)^+.$$

§3. 极限定理证明的特征函数法

1. 证明, 定理 1 中的结论在 $\mathbb{R}^n$ $(n \geqslant 2)$ 时仍然成立.

提示: 证明与一维情形类似, 但是需要建立多维场合下的引理 3, 其形式为:

$$\int_A dF(x) \leqslant \frac{K}{a^n} \int_B (1 - \mathrm{Re}\,\varphi(t))dt,$$

其中

$$A = \left\{x \in \mathbb{R}^n : \ |x_1| \geqslant \frac{1}{a}, \cdots, |x_n| \geqslant \frac{1}{a}\right\},$$
$$B = \{t \in \mathbb{R}^n : \ 0 \leqslant t_1 \leqslant a, \cdots, 0 \leqslant t_n \leqslant a\}.$$

2. (关于大数定律)(a) 设 $\xi_1, \xi_2, \cdots$ 为独立随机变量序列, 具有有限的期望 $\mathbf{E}|\xi_n|$ 与方差 $\mathbf{D}\xi_n \leqslant K$, $n \geqslant 1$. 证明, 大数定律成立, 即对任何 $\varepsilon > 0$, 都有

$$\mathbf{P}\left\{\left|\frac{\xi_1 + \cdots + \xi_n}{n} - \frac{\mathbf{E}(\xi_1 + \cdots + \xi_n)}{n}\right| \geqslant \varepsilon\right\} \to 0, \quad n \to \infty. \qquad (*)$$

(b) 设 $\xi_1, \xi_2, \cdots$ 为具有有限的期望 $\mathbf{E}|\xi_n|$ 与方差 $\mathbf{D}\xi_n \leqslant K$, $n \geqslant 1$ 的随机变量序列, 并且其协方差为 $\mathbf{cov}\,(\xi_i, \xi_j) \leqslant 0$, $i \neq j$. 证明, 大数定律 (*) 成立.

提示: 在 (a) 与 (b) 的证明中应当运用切比雪夫不等式.

(c) (伯恩斯坦) 设 $\xi_1, \xi_2, \cdots$ 为具有有限的期望 $\mathbf{E}|\xi_n|$ 与方差 $\mathbf{D}\xi_n \leqslant K$, $n \geqslant 1$ 的随机变量序列, 并且当 $|i-j| \to \infty$ 时, 其协方差具有性质 $\mathbf{cov}\,(\xi_i, \xi_j) \to 0$. 证明, 此时亦有大数定律 (*) 成立.

提示: 证明, 在所述的条件之下, 有 $\mathbf{D}\frac{\xi_1 + \cdots + \xi_n}{n} \to 0$, $n \to \infty$.

(d) 设 $\xi_1, \xi_2, \cdots$ 为独立同分布的随机变量序列, 令 $\mu_n = \mathbf{E}[\xi_1 I(|\xi_1| \leqslant n)]$. 假设

$$\lim_{x \to \infty} x\mathbf{P}\{|\xi_1| > x\} = 0.$$

证明, 此时有如下形式的大数定律成立:

$$\frac{S_n}{n} - \mu_n \xrightarrow{\mathbf{P}} 0,$$

其中, 如同往常, 有 $S_n = \xi_1 + \cdots + \xi_n$(亦可参阅第 20 题).

提示: 对固定的 $s > 0$, 令 $\xi_i^{(s)} = \xi_i I(|\xi_i| \leqslant s)$ 与 $m_n^{(s)} = \mathbf{E}[\xi_1^{(s)} + \cdots + \xi_n^{(s)}]$. 写

$$\begin{aligned}&\mathbf{P}\{|\xi_1 + \cdots + \xi_n - m_n^{(s)}| > t\} \\ &\quad \leqslant t^{-2}\mathbf{D}(\xi_1^{(s)} + \cdots + \xi_n^{(s)}) + \mathbf{P}\{\xi_1 + \cdots + \xi_n \neq \xi_1^{(s)} + \cdots + \xi_n^{(s)}\}.\end{aligned}$$

利用这一估计式, 推得 (令 $s = n$ 和 $t = \varepsilon n$, $\varepsilon > 0$)

$$\begin{aligned}&\mathbf{P}\left\{\left|\frac{\xi_1 + \cdots + \xi_n}{n} - \mathbf{E}[\xi_1 I(|\xi_1| \leqslant n)]\right| \varepsilon\right\} \\ &\qquad\qquad \leqslant \frac{2}{\varepsilon^2 n}\int_0^n x\mathbf{P}\{|\xi_1| > x\}dx + n\mathbf{P}\{|\xi_1| > n\},\end{aligned}$$

由此即可得出所需的结论.

3. 作为定理 1 的推论, 证明, 函数族 $\{\varphi_n,\ n \geqslant 1\}$ 是等度连续的, 而在每个有限区间上, 收敛性 $\varphi_n \to \varphi$ 都是一致的.

提示: 所谓函数族 $\{\varphi_n,\ n \geqslant 1\}$ 等度连续, 其含义是, 对每个 $\varepsilon > 0$, 都存在 $\delta > 0$, 使得对每个 $n \geqslant 1$, 只要 $|s-t| < \delta$, 就都有 $|\varphi_n(s) - \varphi_n(t)| < \varepsilon$.

如果 $F_n \xrightarrow{w} F$, 则根据第 2 节中的普罗霍洛夫定理, 对一切 $\varepsilon > 0$, 都可找到 $a > 0$, 使得 $\int\limits_{|x| \geqslant a} dF_n < \varepsilon$, $n \geqslant 1$. 如此一来, 就有

$$|\varphi_n(t+h) - \varphi_n(t)| \leqslant \int\limits_{|x| \leqslant a} |e^{itx} - 1|dF_n + 2\varepsilon,$$

由此不难推出函数族 $\{\varphi_n,\ n \geqslant 1\}$ 的等度连续性. 然后利用这种性质, 即可推出, 在每个有限区间 $[a,b]$ 上, 都有

$$\sup_{t \in [a,b]} |\varphi_n(t) - \varphi(t)| \to 0, \quad n \to \infty.$$

4. 设 ξ_n, $n \geqslant 1$, 为随机变量序列, 相应的特征函数序列为 $\varphi_{\xi_n}(t)$, $n \geqslant 1$. 证明, $\xi_n \xrightarrow{d} 0$, 当且仅当, 在 $t=0$ 的某个邻域中, 有 $\varphi_{\xi_n}(t) \to 1$, $n \to \infty$.

提示: 在充分性的证明中应当首先运用引理 3, 由其可以推出测度族 $\{\mathrm{Law}\,(\xi_n),\ n \geqslant 1\}$ 是稠密的.

5. 设 $X_1, X_2, \cdots$ 为独立同分布的 (取值于 $\mathbb{R}^k$ 的) 随机向量序列, 具有零均值和 (有限的) 协方差阵 Γ. 证明

$$\frac{X_1+\cdots+X_n}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, \Gamma).$$

(试比较定理 3.)

提示: 根据第 1 题, 只需对每个 $t \in \mathbb{R}^k$, 证明

$$\mathbf{E}e^{i(t,\xi_n)} \to \mathbf{E}e^{i(t,\xi)}, \quad n \to \infty,$$

其中 $\xi_n = n^{-1/2}(X_1+\cdots+X_n)$, 而 $\xi \sim \mathcal{N}(0, \Gamma)$.

6. 设 $\xi_1, \xi_2, \cdots$ 与 $\eta_1, \eta_2, \cdots$ 为两个随机变量序列, 对每个 n, 随机变量 ξ_n 都与 η_n 独立. 假设当 $n \to \infty$ 时, 有 $\xi_n \xrightarrow{d} \xi$, $\eta_n \xrightarrow{d} \eta$, 其中随机变量 ξ 与 η 相互独立.

(a) 证明, 二维随机变量序列 (ξ_n, η_n) 依分布收敛到 (ξ, η).

(b) 设 $f = f(x, y)$ 为连续函数. 证明, 随机变量 $f(\xi_n, \eta_n)$ 依分布收敛到 $f(\xi, \eta)$.

提示: 收敛性 $(\xi_n, \eta_n) \xrightarrow{d} (\xi, \eta)$ 可以由第 1 题中的命题推出. 为证收敛性 $f(\xi_n, \eta_n) \xrightarrow{d} f(\xi, \eta)$, 需要考察复合函数 $\varphi \circ f:\ \mathbb{R}^2 \to \mathbb{R}$, 其中 $\varphi:\ \mathbb{R} \to \mathbb{R}$ 为有界连续函数.

7. 试举例说明, 定理 2 断言 2 中的关于极限特征函数 $\varphi(t) = \lim\limits_n \varphi_n(t)$ 在 $t=0$ 处的连续性条件, 一般来说, 是不能削弱的. (换言之, 如果 $\varphi(t)$ 不在 $t=0$ 处连续, 则可能出现这样的情况: 尽管 $\varphi_n(t) \to \varphi(t)$, 但是对任何分布函数 F, 却都有 $F_n \stackrel{w}{\nrightarrow} F$.) 试举例说明, 取消 $\varphi(t)$ 在 $t=0$ 处的连续性条件, 可能会破坏与特征函数族 $\varphi_n(t)$, $n \geqslant 1$, 相对应的概率分布族 $\{P_n:\ n \geqslant 1\}$ 的稠密性.

提示: 可以考察期望为 0, 方差为 n 的高斯分布 F_n.

8. 作为引理 3 中的不等式 (4) 的补充, 证明, 如果 ξ 为具有特征函数 $\varphi(t)$ 的随机变量, 则:

(a) 对任何 $a > 0$, 都有

$$\mathbf{P}\{|\xi| \leqslant a^{-1}\} \leqslant \frac{2}{a}\int_{|t|\leqslant a} |\varphi(t)|dt.$$

(b) 对正数 b 和 δ, 有

$$\mathbf{P}\{|\xi| \geqslant b\} \leqslant \frac{\left(1+\dfrac{2\pi}{b\delta}\right)^2}{\delta}\int_0^\delta [1-\mathrm{Re}\ \varphi(t)]dt.$$

(c) 如果 ξ 为非负随机变量, 而 $\psi(a) = \mathbf{E}e^{-a\xi}$ 为其拉普拉斯变换 $(a \geqslant 0)$, 则有

$$\mathbf{P}\{\xi \geqslant a^{-1}\} \leqslant 2(1-\psi(a)).$$

9. 设 $\xi, \xi_1, \xi_2, \cdots$ 为取值于 $\mathbb{R}^n$ 的随机向量. 证明, $\xi_k \xrightarrow{d} \xi$, $k \to \infty$, 当且仅当, 对任何 $t \in \mathbb{R}^n$, 都有

$$(\xi_k,\ t) \xrightarrow{d} (\xi,\ t).$$

(本结果建立在克拉默–沃尔德 (Cramer-Wold) 方法的基础之上, 该方法将验证 $\mathbb{R}^n$ 空间中的随机向量的依分布收敛归结为验证相应的一维随机变量的依分布收敛性.)

10. 作为定理 2 中的辛钦大数律(或称为辛钦准则) 的补充, 试证明如下形式的大数定律:

设 $\xi_1, \xi_2, \cdots$ 为独立同分布随机变量序列, $S_n = \xi_1+\cdots+\xi_n$, $0 < p < 2$. 则对某个常数 $c \in \mathbb{R}$, 有

$$n^{-1/p}S_n \xrightarrow{\mathbf{P}} c,$$

当且仅当, 当 $r \to \infty$ 时①, 有

(a) 当 $p < 1$ 时, $r^p\mathbf{P}\{|\xi_1| > r\} \to 0$, 且 $c = 0$.

(b) 当 $p = 1$ 时, $r\mathbf{P}\{|\xi_1| > r\} \to 0$, 且 $\mathbf{E}[\xi_1 I(|\xi_1| \leqslant r)] \to c$.

(c) 当 $p > 1$ 时, $r^p\mathbf{P}\{|\xi_1| > r\} \to 0$, 且 $\mathbf{E}\xi_1 = c = 0$.

11. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布随机变量序列, $S_n = \xi_1+\cdots+\xi_n$, $n \geqslant 1$. 证明, 当 $n \to \infty$ 时, $n^{-1/2}S_n$ 依概率收敛, 当且仅当, $\mathbf{P}\{\xi_1 = 0\} = 1$.

12. 设 $F(x)$ 与 $(F_n(x))_{n\geqslant1}$ 为分布函数, 而 $\varphi(t)$ 与 $(\varphi_n(t))_{n\geqslant1}$ 为对应的特征函数. 证明, 若有

$$\sup_t |\varphi_n(t) - \varphi(t)| \to 0,$$

则亦有

$$\sup_x |F_n(x) - F(x)| \to 0.$$

13. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布随机变量序列, 服从共同的分布 $F = F(x)$, 而 $S_n = \xi_1+\cdots+\xi_n$, $n \geqslant 1$.

证明大数定律的如下形式 (柯尔莫戈洛夫): 为了存在常数序列 $(a_n)_{n\geqslant1}$, 使得

$$\frac{S_n}{n} - a_n \xrightarrow{\mathbf{P}} 0, \quad n \to \infty, \tag{$*$}$$

必须且只需

$$n\mathbf{P}\{|\xi_1| > n\} \to 0, \quad n \to \infty, \tag{$**$}$$

①原书错为 "当 $r \to 0$ 时" —— 译者注.

或者, 等价地有

$$x[1-F(x)-F(-x)]\to 0,\quad x\to\infty.$$

在这些假设之下, 有 $a_n-\mathbf{E}[\xi_1 I(|\xi_1|\leqslant n)]\to 0,\quad n\to\infty$. (满足条件的数列 $(a_n)_{n\geqslant 1}$ 的存在性称为序列 $\left(\frac{S_n}{n}\right)_{n\geqslant 1}$ 的柯尔莫戈洛夫意义下的稳定性.)

14. 证明, 如果 $\mathbf{E}|\xi_1|<\infty$, 则在上题中有条件 (**) 成立, 此时可取 $a_n\equiv m$, 其中 $m=\mathbf{E}\xi_1$(试比较定理 2, 即关于大数定律成立的辛钦准则, 以及第 10 题和第 13 题中的断言).

15. 设 $\xi_1,\xi_2,\cdots$ 为独立同分布随机变量序列, 取值 $\pm 3,\pm 4,\cdots$ 的概率分别为

$$\mathbf{P}\{\xi_1=\pm x\}=\frac{c}{2x^2\ln x},\quad x=3,4,\cdots,$$

其中正则化常数

$$c=\left(\sum_{k=3}^{\infty}\frac{1}{x^2\ln x}\right)^{-1}.$$

证明, 尽管此处有 $\mathbf{E}|\xi_1|=\infty$, 但是却仍然有第 13 题中的条件 (**) 成立, 并且可以取 $a_n\equiv 0$, 因而意味着 $\frac{S_n}{n}\xrightarrow{\mathbf{P}}0$.

注: 应当指出, 在上述场合下, 由于随机变量 $\xi_1,\xi_2,\cdots$ 不存在有限的数学期望 ($\mathbf{E}|\xi_i|=\infty$), 因此也就不可能成立辛钦形式的大数定律 ($n^{-1}S_n\to m$, 其中 $m=\mathbf{E}\xi_1$; 见定理 2). 但是这并未妨碍其中的随机变量序列 $\xi_1,\xi_2,\cdots$ 具有柯尔莫戈洛夫意义下的稳定性 (参阅第 13 题):

$$\frac{S_n}{n}\xrightarrow{\mathbf{P}}\widetilde{m}\ \ (=0),$$

其中 $\widetilde{m}=\widetilde{\mathbf{E}}\xi_1$ 是由柯尔莫戈洛夫所定义的广义数学期望(见参考文献 [60], 第四章, 第 4 节), 其中

$$\widetilde{\mathbf{E}}\xi_1=\lim_{n\to\infty}\mathbf{E}\Big(I(|\xi_1|\leqslant n)\xi_1\Big).$$

后来, 这种广义数学期望被称为柯尔莫戈洛夫 A-积分. (现在, 在分析学中, 称博雷尔函数 $f=f(x)$, $x\in\mathbb{R}$, 是 *A-可积的*, 如果该函数:

(i) 在弱意义下属于 L^1(亦即 $\lim\limits_n n\,\lambda\{x:|f(x)|>n\}\to 0$),

(ii) 存在极限 $\lim\limits_n\int_{\{x:|f(x)|\leqslant n\}}f(x)\lambda(dx)$, 其中 λ 是 $(\mathbb{R},\mathscr{B}(\mathbb{R}))$ 上的勒贝格测度.

通常将这样定义的积分表示为 (A) $\int f(x)\lambda(dx)$. 有趣的是, 对于这样定义的积分, 许多通常的勒贝格积分的性质, 一般来说, 都是不成立的. 例如, 可加性就可能遭到破坏.)

16. 设 $\xi_1,\xi_2,\cdots$ 为 (具有有限数学期望的) 独立随机变量序列, 对某 $\delta\in(0,1)$, 有

$$\frac{1}{n^{1+\delta}}\sum_{i=1}^{n}\mathbf{E}|\xi_i|^{1+\delta}\to 0,\quad n\to\infty.$$

证明, 该 "$(1+\delta)$ 条件" 可以保证大数定律成立:

$$\frac{1}{n}\sum_{i=1}^{n}(\xi_i-\mathbf{E}\xi_i)\xrightarrow{\mathbf{P}}0,\quad n\to\infty.$$

17. (针对定理 3; 不同分布随机变量场合) 在定理 3 中, 曾经 (运用特征函数方法) 在独立同分布随机变量的场合下给出了中心极限定理的证明, 其中的依据是连续性定理 (定理 1). 运用这一方法也可以证明独立不同分布随机变量场合下的中心极限定理, 其基本思路如下.

设 $A_1,A_2,\cdots$ 为相互独立的随机事件序列, 有 $\mathbf{P}(A_n)=1/n$ (第二章第 4 节第 21 题中的随机事件序列就是该类事件列的例子). 令 $\xi_n=I_{A_n}$, 记 $S_n=\xi_1+\cdots+\xi_n$. 证明

$$\mathbf{E}S_n=\sum_{k\leqslant n}\frac{1}{k}\quad(\sim\ln n,\ \ n\to\infty),$$

$$\mathbf{D}S_n=\sum_{k\leqslant n}\frac{1}{k}\left(1-\frac{1}{k}\right)\quad(\sim\ln n,\ \ n\to\infty).$$

再考察随机变量 $\frac{S_n-\mathbf{E}S_n}{\sqrt{\mathbf{D}S_n}}$ $(n\geqslant 1)$ 的特征函数 $\varphi_n(t)$, 可证明有 $\varphi_n(t)\to e^{-t^2/2}$, 因而 (定理 1) 就有中心极限定理成立: 当 $n\to\infty$ 时, 有

$$\mathbf{P}\left\{\frac{S_n-\mathbf{E}S_n}{\sqrt{\mathbf{D}S_n}}\leqslant x\right\}\to\Phi(x),\quad x\in\mathbb{R}.$$

18. 作为引理 3 中的不等式 (4) 的补充, 证明如下不等式: 对一切 $a>0$, 都有

$$(1-\sin 1)\int_{|x|\geqslant 1/a}dF(x)\leqslant\frac{1}{a}\int_0^a[1-\mathrm{Re}\ \varphi(t)]dt\leqslant 2\int_{|x|\geqslant\sqrt{1/a}}dF(x)+\frac{a}{2}.$$

提示: 为证不等式的右半部, 应将 $1-\mathrm{Re}\ \varphi(t)$ 表示为

$$\begin{aligned}1-\mathrm{Re}\ \varphi(t)&=\int_{\mathbb{R}}(1-\cos tx)dF(x)\\&=\int_{|x|\geqslant\sqrt{1/a}}(1-\cos tx)dF(x)+\int_{|x|<\sqrt{1/a}}(1-\cos tx)dF(x),\end{aligned}$$

分别估计上述两个积分, 然后 (像引理 3 中那样) 利用富比尼定理.

19. 设 $(\xi_n)_{n\geqslant 1}$ 为独立同分布的随机变量序列, 服从柯西分布, 密度函数为

$$\frac{\theta}{\pi(\theta^2+x^2)}, \quad \theta>0, \quad x\in\mathbb{R}.$$

证明, 随机变量 $\frac{1}{n}\max\limits_{i\leqslant n}\xi_i$ 的分布函数 F_n 弱收敛到参数为 $\alpha=1$ 的弗雷歇分布 (参阅第二章第 8 节第 48 题), 即随机变量 $1/T_c$ 的分布, 其中 T_c 服从参数为 $c=\theta/\pi$ 的指数分布:

$$\mathbf{P}\left\{\frac{1}{T_c}\leqslant x\right\}=e^{-c/x}, \quad x>0.$$

20. (离散变量场合下的连续性定理) 设 $\xi,\xi_1,\xi_2,\cdots$ 为取值 $k=0,1,2,\cdots$ 的随机变量序列. 记

$$G(s)=\sum_{k=0}^{\infty}\mathbf{P}\{\xi=k\}s^k \quad 与 \quad G_n(s)=\sum_{k=0}^{\infty}\mathbf{P}\{\xi_n=k\}s^k$$

分别为随机变量 ξ 与 ξ_n $(n\geqslant 1)$ 的母函数.

证明,

$$\lim_n \mathbf{P}\{\xi_n=k\}=\mathbf{P}\{\xi=k\}, \quad k=0,1,2,\cdots,$$

当且仅当

$$\lim_n G_n(s)=G(s), \quad s\in[0,1).$$

21. 试用特征函数方法证明第二章第 10 节第 35 题中的命题. (服从区间 $[-1,1]$ 上的均匀分布的随机变量 U 的特征函数为 $\frac{\sin t}{t}$.)

§4. 独立随机变量之和的中心极限定理 I. 林德伯格条件

1. 设 $\xi_1,\xi_2,\cdots$ 为独立同分布随机变量序列, 有 $\mathbf{E}\xi_1^2<\infty$. 证明

$$\frac{\max(|\xi_1|,\cdots,|\xi_n|)}{\sqrt{n}}\xrightarrow{\text{d}}0, \quad n\to\infty.$$

(试比较第二章第 10 节第 53 题.)

提示: 利用

$$\mathbf{P}\left\{\frac{\max(|\xi_1|,\cdots,|\xi_n|)}{\sqrt{n}}\leqslant\varepsilon\right\}=\left[\mathbf{P}\{\xi_1^2\leqslant n\varepsilon^2\}\right]^n$$

及 $n\varepsilon^2\mathbf{P}\{\xi_1>n\varepsilon\}\to 0$, $n\to\infty$.

2. 试直接证明, 在伯努利场合下 $\sup\limits_x|F_{T_n}(x)-\Phi(x)|$ 的阶为 $\frac{1}{\sqrt{n}}$, $n\to\infty$.①

① $F_{T_n}(x)$ 是随机变量 $T_n=\frac{\xi_1+\cdots+\xi_n}{\sqrt{n}}$ 的分布函数, 其中 $\xi_1,\xi_2,\cdots$ 为独立同分布的对称伯努利随机变量序列, 即有 $\mathbf{P}\{\xi_1=\pm 1\}=1/2$ —— 译者注.

3. 设 $X_1,X_2,\cdots$ 是可交换随机变量的无限序列 (参阅第二章第 5 节第 4 题), 有 $\mathbf{E}X_n=0$, $\mathbf{E}X_n^2=1$, $n\geqslant 1$. 假设

$$\mathbf{cov}\,(X_1,\ X_2)=\mathbf{cov}\,(X_1^2,\ X_2^2). \tag{$*$}$$

证明, 对于这样的随机变量序列, 有中心极限定理成立:

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}X_i\xrightarrow{\text{d}}\mathcal{N}(0,1). \tag{$**$}$$

反之, 若 $(**)$ 成立, 则 $(*)$ 亦成立.

4. (a) 关于格子点随机变量的局部极限定理. 设 $\xi_1,\xi_2,\cdots$ 为独立同分布的随机变量, 有 $\mu=\mathbf{E}\xi_1$, $\sigma^2=\mathbf{D}\xi_1$. 令 $S_n=\xi_1+\cdots+\xi_n$, $n\geqslant 1$, 并假设 $\xi_1,\xi_2,\cdots$ 为格子点的随机变量, 其可能的取值为 $a+kh$, 其中 $k=0,\pm1,\pm2,\cdots$, 且 (最大) 步长为 $h>0$.

证明, 当 $n\to\infty$ 时, 有

$$\sup_k\left|\frac{\sqrt{n}}{h}\mathbf{P}\{S_n=an+kh\}-\frac{1}{\sqrt{2\pi}\,\sigma}\exp\left\{-\frac{(hk+an-n\mu)^2}{2\sigma^2 n}\right\}\right|\to 0.$$

(试比较第二章第 6 节中的局部极限定理.)

提示: 可以运用下面的以自身的特征函数的性质为基础的变换来证明题中结论. 根据第二章第 12 节第 9 题, 有

$$\mathbf{P}\{S_n=an+kh\}=\mathbf{P}\{h^{-1}(S_n-an)=k\}=\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{-iuk}e^{\frac{iuna}{h}}\left[\varphi\left(\frac{u}{h}\right)\right]^n du,$$

其中 $\varphi(u)$ 是随机变量 ξ_1 的特征函数. 另一方面, 显然有

$$e^{-z^2/2}=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{iuz}e^{-u^2/2}du.$$

因此

$$\begin{aligned}&2\pi\left|\frac{\sqrt{n}\sigma}{h}\mathbf{P}\{S_n=an+kh\}-\frac{1}{2\pi}e^{-z^2/2}\right|\\&\leqslant\int_{|t|\leqslant\frac{\pi\sqrt{n}\sigma}{h}}\left|\left[\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n-e^{-t^2/2}\right|dt+\int_{|t|>\frac{\pi\sqrt{n}\sigma}{h}}e^{-t^2/2}dt.\end{aligned}$$

上述表达式的右端与 k 无关, 故只需证明当 $n\to\infty$ 时, 右端趋于 0 即可.

(b) 关于有密度的随机变量的局部极限定理. 设 $\xi_1,\xi_2,\cdots$ 为独立同分布的随机变量, 有 $\mu=\mathbf{E}\xi_1$, $\sigma^2=\mathbf{D}\xi_1$. 假设 ξ_1 的特征函数 $\varphi=\varphi(t)$ 可积, 从而它具有密度函数

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-itx}\varphi(t)dt.$$

(参阅第二章第 12 节定理 3.)

以 $f_n = f_n(x)$ 表示 $S_n = \xi_1 + \cdots + \xi_n\ (n \geqslant 1)$ 的密度函数. 证明, 当 $n \to \infty$ 时, 有

$$\sup_x \left| \sqrt{n} f_n(x) - \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{ -\frac{(x-n\mu)^2}{2\sigma^2 n} \right\} \right| \to 0.$$

提示: 按照格子点情形下的类似步骤进行证明.

5. 设 $X_1, X_2, \cdots$ 为独立同分布的随机变量序列, 有 $\mathbf{E}X_1 = 0$, $\mathbf{E}X_1^2 = 1$. 设 $d_1, d_2, \cdots$ 为非负常数, 有 $d_n = o(D_n)$, 其中 $D_n^2 = \sum_{k=1}^{n} d_k^2$. 证明, "加权" 随机变量序列 $d_1X_1, d_2X_2, \cdots$ 满足中心极限定理:

$$\frac{1}{D_n} \sum_{k=1}^{n} d_k X_k \xrightarrow{\mathrm{d}} \mathcal{N}(0,1).$$

6. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布的随机变量, 有 $\mathbf{E}\xi_1 = 0$, $\mathbf{E}\xi_1^2 = 1$, 令 $S_n = \xi_1 + \cdots + \xi_n$, $n \geqslant 1$. 设 $(\tau_n)_{n\geqslant 1}$ 为仅在集合 $\{1, 2, \cdots\}$ 中取值的随机变量序列, 有 $\tau_n / n \xrightarrow{\mathbf{P}} c$, 其中 $c > 0$ 为常数. 证明,

$$\mathrm{Law}\ (\tau_n^{-1/2} S_{\tau_n}) \to \Phi$$

(亦即 $\tau_n^{-1/2} S_{\tau_n} \xrightarrow{\mathrm{d}} \xi$, 其中 $\xi \sim \mathcal{N}(0,1)$). (应当指出, 此处并未假定序列 $(\tau_n)_{n\geqslant 1}$ 与序列 $(\xi_n)_{n\geqslant 1}$ 的相互独立性.)

7. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布的随机变量, 有 $\mathbf{E}\xi_1 = 0$, $\mathbf{E}\xi_1^2 = 1$. 证明

$$\mathrm{Law}\ (n^{-1/2} \max_{1\leqslant m\leqslant n} S_m) \to \mathrm{Law}\ (|\xi|), \quad \text{其中 } \xi \sim \mathcal{N}(0,1).$$

换言之, 对任何 $x > 0$, 都有

$$\mathbf{P}\left\{ n^{-1/2} \max_{1\leqslant m\leqslant n} S_m \leqslant x \right\} \to \sqrt{\frac{2}{\pi}} \int_0^x e^{-y^2/2} dy \quad \left(= \frac{1}{\sqrt{2}} \mathrm{erf}(x) \right).$$

提示: 首先证明所述结论对相互独立的对称伯努利随机变量序列 $\xi_1, \xi_2, \cdots$ 成立, 其中 $\mathbf{P}\{\xi_n = \pm 1\} = 1/2$. 然后利用所证的结果 (也可直接证明, 但难度较大), 证明, 极限分布的形式对于任何具有所述的矩性质的独立同分布随机变量序列 $\xi_1, \xi_2, \cdots$ 都彼此相同. (这种极限分布与具有 $\mathbf{E}\xi_n = 0$, $\mathbf{E}\xi_n^2 = 1$ 的独立同分布的随机变量序列 $\xi_1, \xi_2, \cdots$ 的具体分布无关的性质称为 "不变原理", 可参阅, 例如, [9] 和 [17].)

8. 在上题中的条件之下, 证明

$$\mathbf{P}\left\{ n^{-1/2} \max_{1\leqslant m\leqslant n} |S_m| \leqslant x \right\} \to H(x), \quad x > 0,$$

其中

$$H(x) = \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} \exp\left\{ -\frac{(2k+1)^2\pi^2}{8x^2} \right\}.$$

9. 设 $X_1, X_2, \cdots$ 为相互独立的随机变量序列, 有

$$\mathbf{P}\{X_n = \pm n^\alpha\} = \frac{1}{2n^\beta}, \quad \mathbf{P}\{X_n = 0\} = 1 - \frac{1}{n^\beta}, \quad \text{其中 } 2\alpha > \beta - 1.$$

证明, 林德伯格条件成立, 当且仅当, $0 \leqslant \beta < 1$.

10. 设 $X_1, X_2, \cdots$ 为独立随机变量序列, 有 $X_n \leqslant C_n$ ($\mathbf{P}$-a.s.), 且 $C_n = o(D_n)$, 其中

$$D_n^2 = \sum_{k=1}^{n} \mathbf{E}(X_k - \mathbf{E}X_k)^2 \to \infty.$$

证明,

$$\frac{S_n - \mathbf{E}S_n}{D_n} \xrightarrow{\mathrm{d}} \mathcal{N}(0,1), \quad \text{其中 } S_n = X_1 + \cdots + X_n.$$

11. 设 $X_1, X_2, \cdots$ 为独立随机变量序列, 有 $\mathbf{E}X_n = 0$, $\mathbf{E}X_n^2 = \sigma_n^2$. 假设对其有中心极限定理成立, 并且

$$\mathbf{E}\left(D_n^{-1/2} \sum_{i=1}^{n} X_i \right)^k \to \frac{(2k)!}{2^k\, k!} \quad \text{对某 } k \geqslant 1.$$

证明, 此时必有k 阶林德伯格条件成立, 亦即

$$\sum_{j=1}^{n} \int_{\{|x|>\varepsilon\}} |x|^k dF_j(x) = o(D_n^k), \quad \varepsilon > 0.$$

(通常的林德伯格条件相当于此处 $k = 2$ 的情形; 参阅 (1) 式.)

12. 设 $X = X(\lambda)$ 与 $Y = Y(\mu)$ 是相互独立的参数分别为 $\lambda > 0$ 和 $\mu > 0$ 的泊松随机变量. 证明,

$$\frac{(X(\lambda) - \lambda) - (Y(\mu) - \mu)}{\sqrt{X(\lambda) + Y(\mu)}} \xrightarrow{\mathrm{d}} \mathcal{N}(0,1), \quad \text{当 } \lambda \to \infty,\ \mu \to \infty.$$

13. 设对每个 $n \geqslant 1$, 随机向量 $(X_1^{(n)}, \cdots, X_{n+1}^{(n)})$ 都是均匀分布在单位球上的 $n+1$ 维的随机向量. 证明, 如下的 (庞加莱) 命题成立: 对一切 $x \in \mathbb{R}$, 都有

$$\lim_{n\to\infty} \mathbf{P}\left\{ \sqrt{n}\, X_{n+1}^{(n)} \leqslant x \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} du.$$

14. 设 $\xi_1, \xi_2, \cdots$ 为相互独立的 $\mathcal{N}(0,1)$ 分布随机变量, 而 $S_0 = 0$, $S_n = \xi_1 + \cdots + \xi_n$, $n \geqslant 1$. 试求 ($n \to \infty$ 时) 如下随机变量的概率分布:

$$\frac{1}{n} \sum_{k=1}^{n} |S_{k-1}| (\xi_k^2 - 1), \quad n \geqslant 1.$$

15. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布的对称伯努利随机变量序列 ($\mathbf{P}\{\xi_1 = 1\} = \mathbf{P}\{\xi_1 = -1\} = 1/2$), 记 $S_0 = 0,\ S_k = \xi_1 + \cdots + \xi_k,\ k \geqslant 1$.

构造连续的过程 $X^{(2n)} = (X_t^{(2n)})_{0 \leqslant t \leqslant 1}$, 其中

$$X_t^{(2n)} = \frac{S_{2nt}}{\sqrt{2n}},$$

而对于 $u \geqslant 0$, 将 S_u 定义为两个相邻的整数时刻之间的线性插值.

试给出证明下列 (重要而困难的) 命题的思路.

(a) $P^{2n} = \mathrm{Law}\,(X_t^{(2n)},\ 0 \leqslant t \leqslant 1)$(在有限维分布弱收敛的意义下, 在赋予了一致距离的距离空间 C 中的分布弱收敛的意义下) 收敛到布朗运动 $B = (B_t)_{0 \leqslant t \leqslant 1}$ 的分布 $P = \mathrm{Law}\,(B_t,\ 0 \leqslant t \leqslant 1)$. ($C$ 中的弱收敛是*唐斯克尔–普罗霍洛夫* (Donsker-Prokhorov) *不变原理*的特殊情况, 所以应当参阅第七章第 8 节第 1 小节.)

(b) 条件分布 $Q^{2n} = \mathrm{Law}\,(X_t^{(2n)},\ 0 \leqslant t \leqslant 1 \mid X_1^{(2n)})$(在 (a) 中所说的意义下) 收敛到布朗桥 $B^{\circ} = (B_t^{\circ})_{0 \leqslant t \leqslant 1}$ 的分布 $Q = \mathrm{Law}\,(B_t^{\circ},\ 0 \leqslant t \leqslant 1)$.

提示: 利用第 13 节中推出柯尔莫戈洛夫极限分布时的分析方法. 更详细的推导可见参考文献 [9] 和 [17].

16. 由上题的结果 (试比较第 7 题与第 8 题中的结果) 推出下列关系式: 对于 $x > 0$,

$$(\mathrm{a}_1)\quad \mathbf{P}\left\{\frac{1}{\sqrt{2n}} \max_{0 \leqslant k \leqslant 2n} S_k \leqslant x\right\} \to \mathbf{P}\left\{\max_{0 \leqslant t \leqslant 1} B_t \leqslant x\right\} \quad (= \mathbf{P}\{|B_1| \leqslant x\}),$$

$$(\mathrm{a}_2)\quad \mathbf{P}\left\{\frac{1}{\sqrt{2n}} \max_{0 \leqslant k \leqslant 2n} |S_k| \leqslant x\right\} \to \mathbf{P}\left\{\max_{0 \leqslant t \leqslant 1} |B_t| \leqslant x\right\}$$

和

$$(\mathrm{b}_1)\ \mathbf{P}\left(\frac{1}{\sqrt{2n}} \max_{0 \leqslant k \leqslant 2n} S_k \leqslant x \,\middle|\, S_{2n} = 0\right) \to \mathbf{P}\left\{\max_{0 \leqslant t \leqslant 1} B_t^{\circ} \leqslant x\right\},$$

$$(\mathrm{b}_2)\ \mathbf{P}\left(\frac{1}{\sqrt{2n}} \max_{0 \leqslant k \leqslant 2n} |S_k| \leqslant x \,\middle|\, S_{2n} = 0\right) \to \mathbf{P}\left\{\max_{0 \leqslant t \leqslant 1} |B_t^{\circ}| \leqslant x\right\}.$$

17. 作为第 15 题和第 16 题的继续, 证明如下各关系式:

$$(\mathrm{a})\ \mathbf{P}\left\{\frac{1}{\sqrt{2n}}\left[\max_{0 \leqslant k \leqslant 2n} S_k - \min_{0 \leqslant k \leqslant 2n} S_k\right] \leqslant x\right\} \to \mathbf{P}\left\{\max_{0 \leqslant t \leqslant 1} B_t - \min_{0 \leqslant t \leqslant 1} B_t \leqslant x\right\}$$

和

$$(\mathrm{b})\ \mathbf{P}\left(\frac{1}{\sqrt{2n}}\left[\max_{0 \leqslant k \leqslant 2n} S_k - \min_{0 \leqslant k \leqslant 2n} S_k\right] \leqslant x \,\middle|\, S_{2n} = 0\right)$$
$$\to \mathbf{P}\left\{\max_{0 \leqslant t \leqslant 1} B_t^{\circ} - \min_{0 \leqslant t \leqslant 1} B_t^{\circ} \leqslant x\right\}.$$

18. 设 $N \in [0, \infty),\ \lambda \in (0, \infty)$. 证明

$$\lim_{n \to \infty} e^{-\lambda n} \sum_{k \leqslant nN} \frac{(\lambda n)^k}{k!} = \begin{cases} 1, & \text{如果 } N > \lambda, \\ 1/2, & \text{如果 } N = \lambda, \\ 0, & \text{如果 } N < \lambda. \end{cases}$$

再证明

$$\lim_{n \to \infty} \left(\sum_{k \leqslant nN} \frac{(\lambda n)^k}{k!}\right)^{1/n} = \begin{cases} e^{\lambda}, & \text{如果 } N \leqslant \lambda, \\ e^{-N \ln \frac{N}{\lambda} + N}, & \text{如果 } N > \lambda. \end{cases}$$

提示: 设 $(\xi_n)_{n \geqslant 1}$ 为相互独立的以 $\lambda > 0$ 为均值的泊松随机变量, 即有 $\mathbf{P}\{\xi_n = k\} = e^{-\lambda}\lambda^k/k!,\ k \geqslant 0$. 验证

$$\mathbf{P}\left\{\frac{\xi_1 + \cdots + \xi_n}{n} \leqslant N\right\} = e^{-\lambda n} \sum_{k \leqslant nN} \frac{(\lambda n)^k}{k!},$$

然后利用中心极限定理的结论.

19. 证明,

$$\frac{1}{n!}\int_0^{n+1} x^n e^{-x} dx \to \frac{1}{2}, \quad n \to \infty.$$

并证明其更为一般的形式:

$$\lim_{n \to \infty} \frac{1}{\Gamma(n+1)} \int_0^{y\sqrt{n+1}+(n+1)} x^n e^{-x} dx = \Phi(y), \quad y \geqslant 0.$$

再证明

$$1 + \frac{n+1}{1!} + \frac{(n+1)^2}{2!} + \cdots + \frac{(n+1)^n}{n!} \sim \frac{1}{2} e^{n+1}, \quad n \to \infty.$$

提示: 利用第二章第 6 节第 80 题和前一题中的结果 (令 $N = \lambda$).

20. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布的随机变量序列, 数学期望 $\mu = \mathbf{E}\xi_1$ 确定, 且有 $\sigma^2 = \mathbf{D}\xi_1 < \infty$. 令 $S_0 = 0,\ S_n = \xi_1 + \cdots + \xi_n$. 证明, 对于最大部分和序列 $M = (M_n)_{n \geqslant 0}$, 其中 $M_n = \max\{S_0, S_1, \cdots, S_n\}$, 有中心极限定理成立: 如果 $0 < \mu < \infty$, 则

$$\lim_{n \to \infty} \mathbf{P}\left\{\frac{M_n - n\mu}{\sigma\sqrt{n}} \leqslant x\right\} = \Phi(x), \quad x \in \mathbb{R}.$$

21. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布的随机变量序列, 有 $\mathbf{E}\xi_1 = 0,\ \mathbf{E}\xi_1^2 = \sigma^2 < \infty$. 令 $S_n = \xi_1 + \cdots + \xi_n,\ n \geqslant 1$. 又设 $\{N_t,\ t \geqslant 0\}$ 是取值于集合 $\{1, 2, \cdots\}$ 的随机变量族, 且当 $t \to \infty$ 时, 有 $N_t/t \xrightarrow{\mathbf{P}} \lambda$, 其中 $0 < \lambda < \infty$.

试证明如下的 (安斯康博 (Ф. Анскомб) 的) 关于中心极限定理成立的结论: 当 $t\to\infty$ 时, 有

$$\mathbf{P}\left\{\frac{S_{N_t}}{\sigma\sqrt{N_t}}\leqslant x\right\}\to\Phi(x),\quad \mathbf{P}\left\{\frac{S_{N_t}}{\sigma\sqrt{\lambda}\sqrt{t}}\leqslant x\right\}\to\Phi(x),$$

其中, $\Phi(x)$ 是标准正态分布的分布函数.

提示: 令 $n_0=[\lambda t]$, 且为简单起见, 设 $\sigma=1$. 将 $S_{N_t}/\sqrt{N_t}$ 表示为

$$\frac{S_{N_t}}{\sqrt{N_t}}=\left(\frac{S_{n_0}}{\sqrt{n_0}}+\frac{S_{N_t}-S_{n_0}}{\sqrt{n_0}}\right)\sqrt{\frac{n_0}{N_t}}.$$

由于 $\mathbf{P}\{S_{n_0}/\sqrt{n_0}\leqslant x\}\to\Phi(x)$, 且当 $t\to\infty$ 时, 有 $n_0/N_t\xrightarrow{\mathbf{P}}1$, 所以只需证明 $(S_{N_t}-S_{n_0})/\sqrt{n_0}\xrightarrow{\mathbf{P}}0$. 为此, 将概率 $\mathbf{P}\{|S_{N_t}-S_{n_0}|>\varepsilon\sqrt{n_0}\}$ 表示为如下的两项之和:

$$\mathbf{P}\{|S_{N_t}-S_{n_0}|>\varepsilon\sqrt{n_0},\ \ N_t\in[n_1,n_2]\}+\mathbf{P}\{|S_{N_t}-S_{n_0}|>\varepsilon\sqrt{n_0},\ \ N_t\notin[n_1,n_2]\},$$

其中 $n_1=[n_0(1-\varepsilon^3)]+1$, $n_2=[n_0(1+\varepsilon^3)]$. 再利用柯尔莫戈洛夫不等式即可断言其趋于 0.

22. (中心极限定理中的矩的收敛性) 设 $\xi_1,\xi_2,\cdots$ 为独立同分布的随机变量序列, 有 $\mathbf{E}\xi_1=0$, $\mathbf{E}\xi_1^2=\sigma^2<\infty$. 令 $S_n=\xi_1+\cdots+\xi_n$, $n\geqslant1$. 根据第 3 节定理 3 或定理 1 中的 b)(参阅第 2 小节), 有

$$\frac{S_n}{\sigma\sqrt{n}}\xrightarrow{\mathrm{d}}N,$$

其中, N 为标准正态分布的随机变量 ($\mathrm{Law}(N)=\mathcal{N}(0,1)$).

证明, 如果对某个 $r\geqslant2$, 有 $\mathbf{E}|\xi_1|^r<\infty$, 则对任何 $0<p\leqslant r$, 都有

$$\mathbf{E}\left|\frac{S_n}{\sigma\sqrt{n}}\right|^p\to\mathbf{E}|\mathrm{N}|^p.$$

提示: 验证随机变量族 $\left\{\left|\frac{S_n}{\sigma\sqrt{n}}\right|^r,\ n\geqslant1\right\}$ 是一致可积的, 再利用第二章第 10 节第 54 题中的断言 b).

23. 设 $(\xi_n)_{n\geqslant1}$ 为独立同分布的随机变量序列, 有 $\xi_n\sim\mathcal{N}(0,1)$. 设 ξ 是与 $(\xi_n)_{n\geqslant1}$ 独立的随机变量, 亦有 $\xi\sim\mathcal{N}(0,1)$.

证明, 极限

$$\lim_{n\to\infty}\mathbf{E}\left|\frac{\xi_1+\cdots+\xi_n}{\sqrt{n}}-\xi\right|$$

存在, 且等于 $2/\sqrt{\pi}$.

提示: 首先验证, 随机变量族 $\{S_n/\sqrt{n}-\xi,\ n\geqslant1\}$ 一致可积, 其中 $S_n=\xi_1+\cdots+\xi_n$.

24. 设 $\mathbf{P}$ 与 $\mathbf{Q}$ 是 $(\Omega,\mathscr{F})$ 上的两个概率测度, 且 $\mathbf{Q}$ 关于 $\mathbf{P}$ 绝对连续 ($\mathbf{Q}\ll\mathbf{P}$). 假设关于测度 $\mathbf{P}$, 序列 $X_1,X_2,\cdots$ 是独立同分布的随机变量序列, 有 $m=\mathbf{E_P}X_i$, $\sigma^2=\mathbf{E_P}(X_i-m)^2$, 其中 $\mathbf{E_P}$ 是按照测度 $\mathbf{P}$ 计算的均值.

此时, 根据第 3 节定理 3, 有中心极限定理成立, 即当 $n\to\infty$ 时, 有

$$\mathbf{P}\left\{\frac{1}{\sigma\sqrt{n}}\sum_{i=1}^n(X_i-m)\leqslant x\right\}\to\Phi(x),\quad x\in\mathbb{R},$$

其中 $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{x}e^{-y^2/2}dy$.

我们回过来看测度 $\mathbf{Q}$. 关于测度 $\mathbf{Q}$, 即使在 $\mathbf{Q}\ll\mathbf{P}$ 的假定下, 一般来说, $X_1,X_2,\cdots$ 未必还是独立同分布的随机变量序列. 但是, 根据雷尼的一个结果, 中心极限定理仍然成立, 即当 $n\to\infty$ 时, 仍然会有

$$\mathbf{Q}\left\{\frac{1}{\sigma\sqrt{n}}\sum_{i=1}^n(X_i-m)\leqslant x\right\}\to\Phi(x),\quad x\in\mathbb{R}.$$

试证之.

提示: 应当证明, 如果 $f=f(x)$ 是有界连续函数, 则

$$\mathbf{E_Q}f(\widehat{S}_n)\to\mathbf{E_P}f(N(0,1)),$$

其中, $\widehat{S}_n=\frac{1}{\sigma\sqrt{n}}\sum\limits_{i=1}^n(X_i-m)$, 而 $N(0,1)$ 是服从标准正态分布 $\mathcal{N}(0,1)$ 的随机变量. 为此, 引入拉东–尼科迪姆导数 $D=\frac{d\mathbf{Q}}{d\mathbf{P}}$ 和随机变量 $D_k=\mathbf{E_P}(D\,|\,\mathscr{F}_k)$, 其中 $\mathscr{F}_k=\sigma(X_1,\cdots,X_k)$, 再把 $\mathbf{E_Q}f(\widehat{S}_n)$ 表示为如下形式:

$$\mathbf{E_Q}f(\widehat{S}_n)=\mathbf{E_P}[(D-D_k)f(\widehat{S}_n)]+\mathbf{E_P}[D_kf(\widehat{S}_n)].$$

然后, 证明 $\lim\limits_k\sup\limits_n|\mathbf{E_P}[(D-D_k)f(\widehat{S}_n)]|=0$; 并证明当 $n\to\infty$ 时, 对任何 $k\geqslant1$, 都有

$$\mathbf{E_P}[D_kf(\widehat{S}_n)]\to\mathbf{E_P}f(N(0,1)).$$

25. 设 $\xi_1,\xi_2,\cdots$ 为独立同分布的随机变量序列, 有

$$\mathbf{P}\{\xi_1>x\}=\mathbf{P}\{\xi_1<-x\},\quad x\in\mathbb{R}\quad 和\quad \mathbf{P}\{|\xi_1|>x\}=x^{-2},\quad x\geqslant1.$$

证明, 当 $n\to\infty$ 时, 有

$$\mathbf{P}\left\{\frac{S_n}{\sqrt{n\ln n}}\leqslant x\right\}\to\Phi(x),\quad x\in\mathbb{R},$$

其中 $S_n=\xi_1+\cdots+\xi_n$, $n\geqslant1$.

注: 这个例子表明, 即使在 $\mathbf{E}\xi_1^2=\infty$ 的情况下, 独立同分布的随机变量序列的部分和 S_n 在适当正则化之后仍有可能渐近于正态分布.

提示: 令 $\xi_{nk}=\xi_k I(|\xi_k|\leqslant\sqrt{n}\ln\ln n)$, 并证明

(i) 当 $n\to\infty$ 时, 有 $\sum_{k=1}^{n}\mathbf{P}\{\xi_{nk}\neq\xi_k\}\to 0$.

(ii) 当 $n\to\infty$ 时, 有 $\mathbf{E}\xi_{nk}^2\sim\ln n$.

(iii) 根据林德伯格定理 (定理 1), 得知 $\frac{1}{\sqrt{n\ln n}}\sum_{k=1}^{n}\xi_{nk}\xrightarrow{\mathrm{d}}\mathcal{N}(0,1)$.

(iv) 当 $n\to\infty$ 时, 有 $\mathbf{P}\left\{S_n\neq\sum_{k=1}^{n}\xi_{nk}\right\}\to 0$.

§5. 独立随机变量之和的中心极限定理 II. 非经典条件

1. 试证明 (5) 式.

提示: 利用

$$\int_{\mathbb{R}}x^2dF_{nk}(x)<\infty \quad 和 \quad \int_{\mathbb{R}}x^2d\Phi_{nk}(x)<\infty$$

证明, (5) 式左右两端的积分都有限. 再利用

$$\begin{aligned}\int_{-\infty}^{\infty}&\left(e^{itx}-itx+\frac{t^2x^2}{2}\right)d(F_{nk}-\Phi_{nk})\\&=\lim_{a\to\infty}\left(e^{itx}-itx+\frac{t^2x^2}{2}\right)(F_{nk}(x)-\Phi_{nk}(x))\Big|_{-a}^{a}\\&\quad-it\int_{-\infty}^{\infty}(e^{itx}-1-itx)[F_{nk}(x)-\Phi_{nk}(x)]dx.\end{aligned}$$

2. 试证明关系式 (10) 和 (12).

3. 设 $N=(N_t)_{t\geqslant0}$ 是第二章第 9 节第 4 小节所引入的更新过程 (即

$$N_t=\sum_{n=1}^{\infty}I(T_n\leqslant t),\ T_n=\sigma_1+\cdots+\sigma_n,$$

其中 $\sigma_1,\sigma_2,\cdots$ 是独立同分布的正值随机变量序列). 令 $\mu=\mathbf{E}\sigma_1<\infty$, $0<\mathbf{D}\sigma_1<\infty$. 证明如下的中心极限定理:

$$\frac{N_t-t\mu^{-1}}{\sqrt{t\mu^{-3}\mathbf{D}\sigma_1}}\xrightarrow{\mathrm{d}}N(0,1),$$

其中, $N(0,1)$ 是标准正态随机变量 (均值为 0, 方差为 1).

§6. 无限可分分布和稳定分布

1. 证明, 如果 $\xi_n\xrightarrow{\mathrm{d}}\xi$, 并且 $\xi_n\xrightarrow{\mathrm{d}}\eta$, 则 $\xi\overset{\mathrm{d}}{=}\eta$.

2. 证明, 如果 φ_1 与 φ_2 都是无穷可分的特征函数, 则 $\varphi_1\cdot\varphi_2$ 也是无穷可分的特征函数.

3. 设 $\varphi_n=\varphi_n(t)$ 为无穷可分的特征函数, 对每个 $t\in\mathbb{R}$, 都有 $\varphi_n(t)\to\varphi(t)$, 其中 $\varphi(t)$ 是某个特征函数. 证明, $\varphi(t)$ 是无穷可分的.

提示: 利用如下事实: 如果 $\varphi_n(t)$ 是无穷可分的, 则存在独立同分布的随机变量 $\xi_1,\cdots,\xi_n$, 使得 $S_n=\xi_1+\cdots+\xi_n$ 的特征函数就是 $\varphi_n(t)$, 并且 $S_n\xrightarrow{\mathrm{d}}T$, 其中 T 是无穷可分的.

4. 证明, 无穷可分分布的特征函数不会等于 0. (参阅第 12 题.)

提示: 固然, 题中结论可以由柯尔莫戈洛夫–莱维–辛钦公式直接推出, 但是我们建议读者不要利用该公式, 而按照如下思路证明之: 如果 φ 是无穷可分的特征函数, 则对于每个 $n\geqslant1$, 存在特征函数 $\varphi_n(t)$, 使得 $\varphi(t)=(\varphi_n(t))^n$, 由此即可断言 $\psi(t)=\lim_{n\to\infty}\psi_n(t)$ 恒等于 1, 其中 $\psi_n(t)=|\varphi_n(t)|^2$.

5. 证明, 伽玛分布是无穷可分分布, 但不是稳定律.

提示: 可以模仿下面的证明: 服从泊松分布的 ξ, 即 $\mathbf{P}\{\xi=k\}=e^{-1}/k!$, 是无穷可分的 (参阅第二章第 8 节第 3 题). 但是它却不是稳定的. 事实上, 如果 $\xi_1+\xi_2\overset{\mathrm{d}}{=}a\xi+b$, 其中 $a>0$, $b\in\mathbb{R}$, 而 ξ_1 与 ξ_2 是 ξ 的相互独立的复制. 那么只能由此得到 $a=1$, $b=0$, 从而就应当有 $\xi_1+\xi_2\overset{\mathrm{d}}{=}\xi$, 而这是不可能的.

6. 证明, 对于 α 稳定的随机变量 ξ, 只要 $r<\alpha$, 就都有 $\mathbf{E}|\xi|^r<\infty$.

提示: 由关于稳定分布的随机变量 ξ 的特征函数 $\varphi(t)$ 的莱维–辛钦表达式可以推出, 存在 $\delta>0$, 使得只要 $t\in(0,\delta)$ 与 $\alpha<2$, 就有 $\mathrm{Re}\,\varphi(t)\geqslant1-c|t|^\alpha$, 其中 $c>0$. 因此, 根据第 3 节引理 3, 对 $a>0$, 有

$$\mathbf{P}\left\{|\xi|\geqslant\frac{1}{a}\right\}\leqslant\frac{cK}{\alpha+1}a^\alpha,$$

此即表明 $\mathbf{P}\{|\xi|^r\geqslant n\}\leqslant\frac{cK}{\alpha+1}n^{-\alpha/r}$, 如果 $r<\alpha$, 那么就有

$$\mathbf{E}|\xi|^r\leqslant1+\sum_{n=1}^{\infty}\mathbf{P}\{|\xi|^r\geqslant n\}<\infty.$$

7. 证明, 如果 ξ 是参数为 $0<\alpha\leqslant1$ 的稳定分布的随机变量, 则其特征函数 $\varphi(t)$ 在 $t=0$ 处不可微.

8. 试直接证明, 如果 $d\geqslant0$, $0<\alpha\leqslant2$, 则函数 $e^{-d|t|^\alpha}$ 是特征函数; 而当 $d>0$, $\alpha>2$ 时, 该函数不是特征函数.

9. 设 $(b_n)_{n\geqslant 1}$ 为数列, 对某 $\delta>0$ 和一切 $|t|<\delta$, 都有 $\lim\limits_n e^{itb_n}$ 存在. 证明 $\overline{\lim\limits_n}|b_n|<\infty$.

提示: 最好按照如下思路进行反证: 假设 $\overline{\lim\limits_n}|b_n|=+\infty$. 于是存在子列, 使得 $\lim\limits_k|b_{n_k}|=\infty$. 为方便起见, 不妨设 $\lim\limits_n|b_n|=\infty$. 记 $h(t)=\lim\limits_n e^{itb_n}$, 其中 $t\in(-\delta,\delta)$. 于是对任何 $[\alpha,\beta]\subseteq(-\delta,\delta)$, 都有

$$\int_{-\delta}^{\delta}I_{[\alpha,\beta]}(t)h(t)dt=\lim_n\int_{-\delta}^{\delta}I_{[\alpha,\beta]}(t)e^{itb_n}dt=0.$$

由此并利用恰当集合原理 (第二章第 2 节), 即可推出, 对任何 $A\in\mathscr{B}([-\delta,\delta])$, 都有 $\int_{-\delta}^{\delta}I_A(t)h(t)dt=0$. 从而, 对任何 $t\in[-\delta,\delta]$, 都有 $h(t)=0$. 另一方面, 由于 $|e^{itb_n}|=1$, 从而对一切 $t\in[-\delta,\delta]$, 都有 $|h(t)|=1$, 导致矛盾, 所以 $\overline{\lim\limits_n}|b_n|<\infty$.

10. 证明, 二项分布、均匀分布和三角形分布都不是无穷可分分布. (将密度函数为 $f(x)=(1-|x|)I_{(-1,1)}(x)$ 的分布称为区间 $(-1,1)$ 上的三角形分布.)

再证明进一步的命题: 任何有限支撑的非退化分布都不是无穷可分分布.

11. 设分布函数 F 与其特征函数 φ 分别可以表示为 $F=F^{(n)}*\cdots*F^{(n)}$ (n 次), $\varphi=[\varphi^{(n)}]^n$, 其中 $F^{(n)}$ 是某个分布函数, 而 $\varphi^{(n)}$ 是它的特征函数, $n\geqslant 1$. 证明, 可以找到一个 ("足够丰富的") 概率空间 $(\Omega,\mathscr{F},\mathbf{P})$, 在它上面能够定义随机变量 T 与 $(\eta_k^n)_{k\leqslant n}$, $n\geqslant 1$ (其中 T 服从分布 F, 而 $\eta_1^n,\cdots,\eta_n^n$ 独立同分布, 均服从分布 $F^{(n)}$), 使得 $T\overset{\mathrm{d}}{=}\eta_1^n+\cdots+\eta_n^n$, $n\geqslant 1$. (亦可参阅《概率》第一卷 p373 中的注.)

12. 试举出一个随机变量的例子, 它不服从无穷可分分布, 而它的特征函数也不会等于 0 (亦可参阅第 4 题).

13. 试证明

(a) 特征函数 $\varphi=\varphi(t)$ 是无穷可分的特征函数, 当且仅当, 对任何 $n\geqslant 1$, n 次方根 $\varphi^{1/n}(t)=e^{\frac{1}{n}\ln\varphi(t)}$ 都是特征函数 (其中 ln 是对数的主值).

(b) 有限个无穷可分的特征函数的乘积是无穷可分的特征函数.

14. 利用上题结论和函数

$$\varphi(t)=\exp\{it\beta+\lambda(e^{ita}-1)\},\quad \lambda>0,\ u\in\mathbb{R},\ \beta\in\mathbb{R}$$

是 (泊松型的) 无穷可分特征函数的事实, 证明, 如下的 (费内提) 函数

$$\varphi(t)=\exp\left\{it\beta+\sum_{j=1}^{k}\lambda_j(e^{itu_j}-1)\right\}$$

与

$$\varphi(t)=\exp\left\{it\beta+\int_{-\infty}^{\infty}(e^{itu}-1)dG(u)\right\}$$

都是无穷可分的特征函数, 其中 $G=G(u)$ 是有界上升函数.

15. $\varphi=\varphi(t)$ 是某个具有有限二阶矩的分布的特征函数. 证明, 此类函数是无穷可分分布的特征函数, 当且仅当, $\varphi(t)$ 具有柯尔莫戈洛夫表达式

$$\varphi(t)=\exp\psi(t),$$

而

$$\psi(t)=itb+\int_{-\infty}^{\infty}(e^{itu}-1-itu)\frac{1}{u^2}dG(u),$$

其中 $b\in\mathbb{R}$, $G=G(u)$ 是非降左连续函数, 有 $G(-\infty)=0$, $G(\infty)<\infty$. (试比较上题中的 (费内提) 函数.)

16. 证明, 如果 $\varphi(t)$ 是无穷可分分布的特征函数, 则对任何 $\lambda\geqslant 0$, 函数 $\varphi^\lambda(t)$ 都是特征函数.

17. (续柯尔莫戈洛夫–莱维–辛钦公式) 设 $h=h(x)$ 是对 $x\in\mathbb{R}$ 有定义的截断函数 (在 $x=0$ 的邻域中满足条件 $h(x)=x$, 具有紧支撑的有界连续函数). 证明,

(a) 柯尔莫戈洛夫–莱维–辛钦公式 (2) 可以表示成如下形式:

$$\varphi(t)=\exp\psi_h(t),$$

而

$$\psi_h(t)=itb-\frac{t^2c}{2}+\int_{-\infty}^{\infty}(e^{itx}-1-ith(x))F(dx),$$

其中 $b=b(h)\in\mathbb{R}$, $c\geqslant 0$, $F(dx)$ 是 $(\mathbb{R},\mathscr{B}(\mathbb{R}))$ 上的测度, 有 $F(\{0\})=0$ 和 $\int(x^2\wedge 1)dF(x)<\infty$.

(b) 对于不同的截断函数 h 和 h', 系数 $b(h)$ 与 $b(h')$ 满足关系式

$$b(h')=b(h)+\int(h'(x)-h(x))F(dx).$$

(c) 如果 $\varphi(t)$ 所对应的分布具有有限的二阶矩, 则 $\int_{-\infty}^{\infty}x^2F(dx)<\infty$.

18. 证明, 密度函数为

$$f(x)=\frac{1}{\sqrt{2\pi x^3}}e^{-1/(2x)},\quad x>0$$

的分布是参数为 $\alpha=\frac{1}{2}$, $\beta=0$, $\theta=-1$ 的稳定分布 (参阅 (10) 式).

19. 称随机变量 ξ_m 服从参数为 $\lambda(\{x_m\})>0$ 的广义泊松分布, 如果

$$\mathbf{P}\{\xi_m=kx_m\}=\frac{e^{-\lambda(\{x_m\})}\lambda^k(\{x_m\})}{k!},\quad k=0,1,2,\cdots,$$

其中 $x_m\in\mathbb{R}\setminus\{0\}$.

设 $\xi_1, \cdots, \xi_n$ 是相互独立的所述类型的随机变量. 以 $\lambda = \lambda(dx)$ 表示 $\mathbb{R}\setminus\{0\}$ 中的以 n 点集合 $\{x_m,\ m = 1, \cdots, n\}$ 为支撑的测度, 它在点 x_m 处的测度值为 $\lambda(\{x_m\})$. 令 $T_n = \xi_1 + \cdots + \xi_n$. 随机变量 T_n 的分布称为*复合泊松分布*.

复合泊松分布随机变量 T_n 的特征函数 $\varphi_{T_n}(t)$ 具有如下的表达式:

$$\varphi_{T_n}(t) = \exp\left\{\int\limits_{\mathbb{R}\setminus\{0\}} (e^{itx} - 1)\lambda(dx)\right\}.$$

(由此表达式看出, 复合泊松分布是无穷可分的, 而它与柯尔莫戈洛夫–莱维–辛钦公式 (2) 的结合展示出了该种分布在整个无穷可分分布类中所起的 "结构" 上的关键作用. 这种作用就是: 每一个无穷可分分布都是某个复合泊松分布序列的 (弱) 极限.)

20. 在概率空间 $(\Omega, \mathscr{F}, \mathbf{P})$ 上考察这样的事件组序列: $A^{(n)} = (A_{nk},\ 1 \leqslant k \leqslant n)$, $n \geqslant 1$, 其中, 对每个 n, 事件 $A_{n1}, \cdots, A_{nn}$ 都相互独立. 假设

$$\lim_{n\to\infty} \max_{1\leqslant k\leqslant n} \mathbf{P}(A_{nk}) = 0$$

和

$$\lim_{n\to\infty} \sum_{k=1}^{n} \mathbf{P}(A_{nk}) = \lambda, \quad 其中\ \lambda > 0.$$

试证明如下的命题 (*稀有事件定律*): 随机变量序列 $\xi^{(n)} = \sum_{k=1}^{n} I(A_{nk})$ 依分布收敛到某个服从参数为 $\lambda > 0$ 的泊松分布的随机变量 ξ.

21. 设 X 与 Y 为相互独立的随机变量, 均服从参数为 $\lambda > 0$ 的泊松分布. 试求随机变量 $X - Y$(有时也称其为*双向泊松分布随机变量*) 的特征函数 $\varphi(t)$. 证明, 随机变量 $X - Y$ 的分布为复合泊松分布 (参阅第 19 题, 亦可参阅第二章第 8 节第 3 题).

22. 设 $\xi^{(n)} = (\xi_{nk},\ 1 \leqslant k \leqslant n)$, $n \geqslant 1$ 是随机变量组序列, 其中, 对每个 n, 随机变量 $\xi_{n1}, \cdots, \xi_{nn}$ 都相互独立. 设 $\varphi_{nk} = \varphi_{nk}(t)$ 是随机变量 ξ_{nk} 的特征函数. 证明, 如下二条件相互等价:

(a) $\lim\limits_{n\to\infty} \max\limits_{1\leqslant k\leqslant n} \mathbf{P}\{|\xi_{nk}| > \varepsilon\} = 0$ (随机变量组序列 $(\xi^{(n)},\ n \geqslant 1)$ 的*渐近无穷小条件或称极限无穷小条件*).

(b) 对每个 $t \in \mathbb{R}$, 都有 $\lim\limits_{n\to\infty} \max\limits_{1\leqslant k\leqslant n} |1 - \varphi_{nk}(t)| = 0$.

23. 我们说, 随机变量 ξ 服从 (参数为 $\rho > 0$, $b > 0$ 的*连续的*) *帕雷托*(Pareto)*分布*, 如果它具有如下形式的密度函数 $f_{\rho,b}(x)$:

$$f_{\rho,b}(x) = \frac{\rho b^{\rho}}{x^{\rho+1}} I(x \geqslant b).$$

证明, 该分布是无穷可分的, 而随机变量 $\log \frac{\xi}{b}$ 服从参数为 ρ 的指数分布.

注: 离散型的帕雷托 (Pareto) 分布见第二章第 8 节第 85 题.

24. 我们说, 在 $(0, \infty)$ 中取值的随机变量 ξ 服从参数为 (μ, ρ) 的*逻辑斯谛分布* (logistic distribution), 其中 $\mu \in \mathbb{R}$, $\rho > 0$, 如果

$$\mathbf{P}\{\xi \leqslant x\} = \frac{1}{1 + e^{-(x-\mu)/\rho}}, \quad x > 0.$$

证明, 该分布是无穷可分的.

§7. 弱收敛的 "可度量性"

1. 证明, 在空间 $E = \mathbb{R}$ 中, 概率分布 P 与 $\widetilde{P}$ 之间的莱维–普罗霍洛夫距离 $L(P, \widetilde{P})$ 不小于与 P 和 $\widetilde{P}$ 相对应的分布函数 F 与 $\widetilde{F}$ 之间的莱维距离 $L(F, \widetilde{F})$(参阅第 1 节第 4 题). 并举出成立严格不等号的例子.

提示: 为证明 $L(F, \widetilde{F}) \leqslant L(P, \widetilde{P})$, 只需证明

$$L(F, \widetilde{F}) = \inf\{\varepsilon > 0:\ P(D) \leqslant \widetilde{P}(D^{\varepsilon}) + \varepsilon\ 和\ \widetilde{P}(D) \leqslant P(D^{\varepsilon}) + \varepsilon\ 对一切形如\ (-\infty, x]\ 的集合\ D,\ x \in \mathbb{R}\},$$

而

$$L(P, \widetilde{P}) = \inf\{\varepsilon > 0:\ P(D) \leqslant \widetilde{P}(D^{\varepsilon}) + \varepsilon\ 和\ \widetilde{P} \leqslant P(D^{\varepsilon}) + \varepsilon\ 对一切闭集\ D \subseteq \mathbb{R}\}.$$

为举出成立严格不等号的例子, 可以考察测度 $P = \delta_0$ 和 $\widetilde{P} = \frac{1}{2}(\delta_{-1} + \delta_1)$, 其中 δ_a 是集中在点 a 处的单点概率测度, 即

$$\delta_a(A) = \begin{cases} 1, & 如果\ a \in A, \\ 0, & 如果\ a \notin A. \end{cases}$$

(在本例中, 有 $L(F, \widetilde{F}) = \frac{1}{2}$, $L(P, \widetilde{P}) = 1$, 试证之.)

2. 证明, (19) 式定义了空间 BL 中的一个距离.

提示: 为证性质 $\|P - \widetilde{P}\|^*_{BL} = 0 \Rightarrow P = \widetilde{P}$ (对距离的其他性质的验证都很简单), 请考虑对于闭集 A 和 $\varepsilon > 0$ 通过 (14) 式所定义的函数 $f_A^{\varepsilon}(x)$. 当 $\varepsilon \downarrow 0$ 时, 有

$$\int f_A^{\varepsilon}(x)P(dx) \to P(A) \quad 和 \quad \int f_A^{\varepsilon}(x)\widetilde{P}(dx) \to \widetilde{P}(A),$$

所以对任何闭集 A, 都有 $P(A) = \widetilde{P}(A)$. 再令 $\mathscr{M} = \{A \in \mathscr{B}(E):\ P(A) = \widetilde{P}(A)\}$, (并利用 "恰当集合原理" 和 "$\pi$-$\lambda$ 系" 的性质, 参阅第二章第 2 节,) 即可断言 $\mathscr{M} = \mathscr{B}(E)$.

3. 证明不等式 (20), (21) 和 (22).

4. 设 $F \doteq F(x)$ 与 $G = G(x)$, $x \in \mathbb{R}$ 是两个分布函数, P_c 与 Q_c 分别是它们与直线 $x + y = c$ 的交点. 证明, F 与 G 的莱维距离 (参阅第 1 节第 4 题) 等于

$$L(F, G) = \sup_c \frac{\overline{P_c Q_c}}{\sqrt{2}},$$

其中 $\overline{P_c Q_c}$ 是连接点 P_c 与 Q_c 的线段的长度.

5. 证明, 赋予莱维距离的全体分布函数的集合是一个波利希 (Polish) 空间 (完备可分的距离空间).

6. 设 $K(F, G)$ 是分布函数 F 与 G 之间的柯尔莫戈洛夫距离:

$$K(F, G) = \sup_x |F(x) - G(x)|,$$

而 $L(F, G)$ 是它们之间的莱维距离. 证明

$$L(F, G) \leqslant K(F, G),$$

而若 G 绝对连续, 则还有

$$K(F, G) \leqslant (1 + \sup_x |G'(x)|) L(F, G).$$

7. 设 X 与 $\widetilde{X}$ 是定义在同一个概率空间上的两个随机变量, 分别具有分布函数 F 与 $\widetilde{F}$. 证明, 它们之间的莱维距离 $L(F, \widetilde{F})$ 满足关系式

$$L(F, \widetilde{F}) \leqslant d + \mathbf{P}\{|X - \widetilde{X}| > d\}, \quad \forall\, d > 0,$$

和

$$L(F, \widetilde{F}) \leqslant (c+1) e^{\frac{c}{c+1}} (\mathbf{E}|X - \widetilde{X}|^c)^{\frac{1}{c+1}}, \quad \forall\, c \geqslant 1.$$

8. 利用刚才的第 6 题与第 7 题的结果, 证明, 如果 X 与 $\widetilde{X}$ 是定义在同一个概率空间上的两个随机变量, 分别具有分布函数 F 与 $\widetilde{F}$, 而 $\Phi(x)$ 是标准正态分布 $\mathscr{N}(0,1)$ 的分布函数, $\sigma > 0$, 则有

$$\sup_x \left| F(x) - \Phi\left(\frac{x}{\sigma}\right) \right| \leqslant \left(1 + \frac{1}{\sqrt{2\pi\sigma^2}}\right) \left[\sup_x \left| \widetilde{F}(x) - \Phi\left(\frac{x}{\sigma}\right) \right| + 2(\mathbf{E}|X - \widetilde{X}|^2)^{1/2} \right].$$

§8. 关于测度的弱收敛与随机元的几乎处处收敛之间的联系

1. 证明, 在可分的距离空间情况下, 对于任何定义在某个概率空间 $(\Omega, \mathscr{F}, \mathbf{P})$ 上的随机元 $X(\omega)$ 与 $Y(\omega)$, 实值函数 $\rho(X(\omega), Y(\omega))$ 都是随机变量.

提示: 设 $\{z_1, z_2, \cdots\}$ 是 E 中的可数稠密子集. 证明, 对任何 $a > 0$, 都有

$$\{\omega:\ \rho(X(\omega), Y(\omega)) < a\} = \bigcap_{n=1}^{\infty} \bigcup_{m=1}^{\infty} \left(\left\{\omega:\ \rho(X(\omega), z_m) < \frac{1}{n}\right\} \cap \left\{\omega:\ \rho(Y(\omega), z_m) < a - \frac{1}{n}\right\} \right),$$

由此并根据第二章第 4 节引理 1, 即可断言 $\rho(X(\omega), Y(\omega))$ 是 $\mathscr{F}$ 可测的.

2. 证明, 由 (2) 式所定义的函数 $d_{\mathbf{P}}(X, Y)$ 是取值于 E 的随机元空间中的一个距离.

提示: 根据上一题, 集合 $\{\rho(X, Y) < \varepsilon\}$ 是可测的, 这就表明, 量 $d_{\mathbf{P}}(X, Y)$ 已经适当地定义下来. 再直接验证距离的三条性质.

3. 证明蕴涵关系式 (5).

4. 证明, 集合 $\Delta_h = \{x \in E:\ h(x)$ 在点 x 处不 ρ 连续$\} \in \mathscr{E}$.

提示: 设 $\{a_1, a_2, \cdots\}$ 是 E 中的可数稠密子集. 为证 $\Delta_n \in \mathscr{E}$, 只需证明如下表达式的正确性:

$$\Delta_n = \bigcup_{n=1}^{\infty} \bigcap_{m=1}^{\infty} \bigcup_{k=1}^{\infty} A_{n,m,k},$$

其中

$$A_{n,m,k} = \begin{cases} B_{1/m}(a_k), & \text{如果存在 } y, z \in B_{1/m}(a_k), \text{ 使得 } |h(y) - h(z)| > 1/n, \\ \varnothing, & \text{其余情况.} \end{cases}$$

易知 $A_{n,m,k}$ 属于 $\mathscr{E}$.

5. 设随机变量对 (ξ, η) 与 $(\widetilde{\xi}, \widetilde{\eta})$ 依分布重合 (即 $(\xi, \eta) \stackrel{\mathrm{d}}{=} (\widetilde{\xi}, \widetilde{\eta})$), 并且 $\mathbf{E}|\xi| < \infty$. 证明, $\mathbf{E}(\xi \,|\, \eta) \stackrel{\mathrm{d}}{=} \mathbf{E}(\widetilde{\xi} \,|\, \widetilde{\eta})$.

6. ξ 与 η 是 (在某个足够丰富的概率空间中给出的) 两个取值于博雷尔空间 $(E, \mathscr{E})$ 的随机元 (参阅第二章第 7 节定义 9). 证明, 可以找到某个定义在 $E \times [0,1]$ 上的取值于 E 的可测函数 $f = f(x, y)$ 和某个服从区间 $[0,1]$ 上的均匀分布的随机变量 α, 使得如下关系式成立:

$$\xi = f(\eta, \alpha).$$

7. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布随机变量序列, 有 $\mathbf{E}\xi_1 = 0$, $\mathbf{E}\xi_1^2 < \infty$. 令 $X_n = \sum_{k=1}^{n} \xi_k$, $n \geqslant 1$. 试证明如下结果 (称为斯科罗霍德勘入): 存在某个概率空间 $(\widetilde{\Omega}, \widetilde{\mathscr{F}}, \widetilde{\mathbf{P}})$, 在其上面定义了一个布朗运动 $\widetilde{B} = (\widetilde{B}_t)_{t \geqslant 0}$ 和一列停时 $\widetilde{\tau} = (\widetilde{\tau}_k)_{k \geqslant 0}$, 有 $0 = \widetilde{\tau}_0 \leqslant \widetilde{\tau}_1 \leqslant \cdots$, 使得

$$(X_n)_{n \geqslant 1} \stackrel{\mathrm{d}}{=} (\widetilde{B}_{\tau_n})_{n \geqslant 1},$$

并且 $\widetilde{\mathbf{E}}(\tau_n - \tau_{n-1}) = \mathbf{E}\xi_1^2$, $n \geqslant 1$. (符号 $\stackrel{\mathrm{d}}{=}$ 意味着同分布.)

8. 设 $F = F(x)$ 是 $\mathbb{R}$ 上的分布函数. 定义其*反函数* $F^{-1}(u)$, $0 \leqslant u \leqslant 1$, 为:

$$F^{-1}(u) = \begin{cases} \inf\{x:\ F(x) > u\}, & \text{如果 } u < 1, \\ \infty, & \text{如果 } u = 1. \end{cases}$$

试证明,

(a) $\{x:\ F(x) > u\} \subseteq \{x:\ F^{-1}(u) \leqslant x\} \subseteq \{x:\ F(x) \geqslant u\}$.

(b) $F(F^{-1}(u)) \geqslant u,\quad F^{-1}(F(x)) \geqslant x$.

(c) 若 $F = F(x)$ 为连续函数, 则有 $F^{-1}(u) = \inf\{x : F(x) \geqslant u\}$, $F^{-1}(u) = \max\{x : F(x) = u\}$, $F(F^{-1}(u)) \geqslant u$ 和 $\{x : F(x) > u\} = \{x : F^{-1}(u) < x\}$.

(d) $\inf\{x : F(x) \geqslant u\} = \sup\{x : F(x) < u\}$.

注: 在数理统计中, 函数 $Q(u) = F^{-1}(u)$ 称为*分位点函数*①.

9. 设 $F = F(x)$ 是分布函数, 而 $F^{-1} = F^{-1}(u)$ 是它的反函数.

(a) 证明, 如果 U 是服从区间 [0, 1] 上的均匀分布的随机变量, 则随机变量 $F^{-1}(U)$ 的分布就是 $F = F(x)$, 亦即

$$\mathbf{P}\{F^{-1}(U) \leqslant x\} = F(x).$$

(b) 证明, 如果随机变量 X 的分布函数 $F = F(x)$ 连续, 则随机变量 $F(X)$ 具有区间 [0, 1] 上的均匀分布.

注: 设 $C(u) = \mathbf{P}\{U \leqslant u\}$ 是区间 [0, 1] 上均匀分布的随机变量 U 的分布函数, 则有 $C(F(x)) = F(x)$. 试比较第 12 题的结果.

10. 设 $F(x, y)$ 是随机变量对 (ξ, η) 的分布函数, 而 $F_1(x) = \mathbf{P}\{\xi \leqslant x\}$, $F_2(y) = \mathbf{P}\{\eta \leqslant y\}$ 分别为变量 ξ 与 η 的分布函数. 证明如下的*弗雷歇–霍夫丁不等式*: 对任何 $x, y \in \mathbb{R}$, 都有

$$\max(F_1(x) + F_2(y) - 1,\ 0) \leqslant F(x, y) \leqslant \min(F_1(x),\ F_2(y)).$$

11. 设 (U, V) 是取值于 $[0, 1]^2$ 的随机向量, 分布函数为

$$C(u, v) = \mathbf{P}\{U \leqslant u,\ V \leqslant v\},$$

并且 U 与 V 都服从区间 [0, 1] 上的均匀分布. 又设 $F_1(x)$ 与 $F_2(y)$ 是两个连续的分布函数, $x, y \in \mathbb{R}$.

证明, 函数

$$F(x, y) = C(F_1(x),\ F_2(y)), \quad x, y \in \mathbb{R}, \tag{$*$}$$

①原书中有如下的一段话: "我们指出,《概率》第一卷第三章第 8 节关于分位点函数的定义式 (4) 中的 $F(x) \geqslant u$ 应当是 $F(x) > u$; (5) 式则应当换为本题中的结论 (a); 参阅本书末尾关于《概率》第一、二卷的勘误表." 在本次出版的中译本中已经按照勘误表作了更正 —— 译者注.

是二元分布函数, 并且它的两个边缘分布为 $F_1(x)$ 和 $F_2(y)$.

注: 人们曾经对这样的问题表现出极大的兴趣: 如何根据给定的具有边缘分布 $F_1(x)$ 和 $F_2(y)$ 的二元分布函数 $F(x, y)\,(x, y \in \mathbb{R})$ 来构造满足性质 $(*)$ 的函数 $C(u, v)$? 现在的这种做法, 即通过某两个取值于区间 [0, 1] 的随机变量 U 与 V 的分布函数 $\mathbf{P}\{U \leqslant u,\ V \leqslant v\}$ 来得到具有所述性质的函数的办法, 是斯克利亚尔 (A. Sklar) 于 1959 年给出的, 称为*系耦* (copula). 在他的著作 (参考文献 [110]) 中, 证明了这种函数的存在性 (和唯一性). (作为例子, 可参阅下一题.)

12. 设 $F(x, y)$ 为二元分布函数, 由下式所定义:

$$F(x, y) = \max(x + y - 1,\ 0),$$

其中 $0 \leqslant x, y \leqslant 1$.

(a) 证明, 它的两个边缘分布 $F_1(x)$ 与 $F_2(y)$ 都是区间 [0, 1] 上的均匀分布.

(b) 证明, 如果*系耦*如下式所定义:

$$C(u, v) = \max(u + v - 1,\ 0), \quad 0 \leqslant u, v \leqslant 1,$$

则有

$$F(x, y) = C(F_1(x),\ F_2(y)).$$

注: 试比较第 9 题的结论.

13. 设 ξ 与 $\xi_1, \xi_2, \cdots$ 为随机变量, 有 $\mathrm{Law}(\xi_n) \to \mathrm{Law}(\xi)$. 假设 $\xi_n \geqslant 0$. 证明

$$\mathbf{E}\,\xi \leqslant \varliminf_n \mathbf{E}\,\xi_n.$$

提示: 利用现在第 8 节中的定理 1 和第二章第 6 节中的定理 2 (法图引理).

§9. 概率测度之间的变差距离. 角谷–海林格距离和海林格积分. 对测度的绝对连续性和奇异性的应用

1. 利用引理 2 的符号, 令

$$P \wedge \widetilde{P} = E_Q(z \wedge \widetilde{z}),$$

其中 $z \wedge \widetilde{z} = \min(z, \widetilde{z})$. 证明

$$\|P - \widetilde{P}\| = 2(1 - P \wedge \widetilde{P}).$$

(因而就有 $\mathscr{E}_r(P, \widetilde{P}) = P \wedge \widetilde{P}$, 关于 $\mathscr{E}_r(P, \widetilde{P})$ 的定义可参阅《概率》第一卷第三章第 9 节中的 (4) 式.)

提示: 利用表达式 $a \wedge b = \frac{1}{2}(a + b - |a - b|)$.

2. 设 P, P_n, $n \geqslant 1$, 是 $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ 上的一列概率测度, 分别具有 (相对于勒贝格测度的) 密度函数 $p(x)$, $p_n(x)$, $n \geqslant 1$. 假设按勒贝格测度对几乎所有的 x, 都有 $p_n(x) \to p(x)$. 证明

$$\|P - P_n\| = \int_{-\infty}^{\infty} |p(x) - p_n(x)|dx \to 0, \quad n \to \infty.$$

提示: 按照如下方式表示所考察的积分式:

$$\int_{-\infty}^{\infty} |p(x) - p_n(x)|dx \leqslant \int_{\{|x|\leqslant a\}} |p(x) - p_n(x)|dx + \int_{\{|x|>a\}} p(x)dx + \int_{\{|x|>a\}} p_n(x)dx,$$

其中 $a > 0$ 应当如此选取, 使得对于任意给定的 $\varepsilon > 0$, 有 $\int_{\{|x|\leqslant a\}} p(x)dx > 1-\varepsilon$. 然后利用法图引理, 得知

$$\varliminf_n \int_{\{|x|\leqslant a\}} p_n(x)dx \geqslant 1 - \varepsilon.$$

3. 设 P 与 $\widetilde{P}$ 是两个概率测度. 将利 P 反 $\widetilde{P}$ 的库尔贝克 (Kullback) 信息 $K(P, \widetilde{P})$ 定义为

$$K(P, \widetilde{P}) = \begin{cases} E \ln \dfrac{dP}{d\widetilde{P}}, & \text{如果 } P \ll \widetilde{P}, \\ \infty, & \text{其余场合.} \end{cases}$$

证明

$$K(P, \widetilde{P}) \geqslant -2\ln(1 - \rho^2(P, \widetilde{P})) \geqslant 2\rho^2(P, \widetilde{P}),$$

其中 $\rho(P, \widetilde{P})$ 是测度 P 与 $\widetilde{P}$ 之间的角谷–海林格 (Kakutani-Helinger) 距离.

提示: 第二个不等号是不等式 $-\ln(1-x) \geqslant x$ 的自然推论, $0 \leqslant x \leqslant 1$. 而为证第一个不等号, 需首先确认

$$-2\ln(1 - \rho^2(P, \widetilde{P})) = -2\ln E_P\sqrt{\frac{\widetilde{z}}{z}},$$

然后利用延森 (Jensen) 不等式, 证明

$$-2\ln E_P\sqrt{\frac{\widetilde{z}}{z}} \leqslant K(P, \widetilde{P}).$$

4. 试证明 (11), (12) 式.

5. 试证明不等式 (24).

提示: 如果 $Q = \frac{1}{2}(P + \widetilde{P})$, $z = \frac{dP}{dQ}$, $\widetilde{z} = \frac{d\widetilde{P}}{dQ}$, 令 $y = z - 1$, 则有 $\widetilde{z} = 2 - z = 1 - y$, 而不等式 (24) 具有形式

$$2(1 + E_Q f(y)) \leqslant 2E_Q|y| \leqslant \sqrt{c_\alpha(1 - E_Q f(y))},$$

其中 $f(y) = (1+y)^\alpha(1-y)^{1-\alpha}$, $y \in [-1, 1]$. 再在区间 $(-1, 1)$ 上面分析 $f'(y)$ 与 $f''(y)$, 推导出:

(a) $f = f(y)$ 是区间 $[-1, 1]$ 中的上凸函数, 且有 $f(y) \geqslant 1 - |y|$.

(b) $f(y) \leqslant 1 + f'(0)y - \widetilde{c}_\alpha y^2$, $y \in [-1, 1]$, 其中 $\widetilde{c}_\alpha = \alpha(1-\alpha)/4$.

然后由 (a) 推出所要证明的第一个不等号, 由 (b) 推出第二个不等号.

6. 设 P, $\widetilde{P}$, Q 是 $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ 上的概率测度, $P * Q$ 与 $\widetilde{P} * Q$ 是它们的卷积 (参阅第二章第 8 节第 4 小节). 证明

$$\|P * Q - \widetilde{P} * Q\| \leqslant \|P - \widetilde{P}\|.$$

提示: 利用引理 1.

7. 试证明例 2 中的性质 (30).

提示: 通过直接计算, 可得

$$H\left(\frac{1}{2};\ P,\ \widetilde{P}\right) = \exp\left\{-\frac{1}{2}\sum_{k=1}^{\infty}\left(\sqrt{\lambda_k} - \sqrt{\widetilde{\lambda}_k}\right)^2\right\}.$$

接下来应当利用定理 2 和定理 3.

8. 设 ξ 与 η 是 $(\Omega, \mathscr{F}, \mathbf{P})$ 上的取值于 $(E, \mathscr{E})$ 的随机元. 证明

$$|\mathbf{P}\{\xi \in A\} - \mathbf{P}\{\eta \in A\}| \leqslant \mathbf{P}(\xi \neq \eta), \quad A \in \mathscr{E}.$$

提示: 利用

$$|I(\xi \in A) - I(\eta \in A)| = |I(\xi \in A) - I(\eta \in A)|I(\xi \neq \eta).$$

9. 如下的公式

$$H(\alpha;\ P,\ \widetilde{P}) = \int_\Omega (dP)^\alpha (d\widetilde{P})^{1-\alpha}$$

(参阅 (20) 式) 定义了测度 P 与 $\widetilde{P}$ 之间的 α 阶海林格积分. 在概率–统计试验的许多问题中, 考察按如下方式定义的所谓海林格变换 $H(\alpha;\ \mathscr{E})$ 似乎更加有益.

令 $\mathscr{E} = (\Omega, \mathscr{F};\ P_0, P_1, \cdots, P_k)$ 是一个概率–统计试验, 意即一个可测空间 $(\Omega, \mathscr{F})$ 连同在它上面所给出的概率测度 $P_0, P_1, \cdots, P_k$.

用符号的形式, 试验 $\mathscr{E}$ 的海林格变换 $H(\alpha;\ \mathscr{E})$ 的定义是:

$$H(\alpha;\ \mathscr{E}) = \int_\Omega (dP_0)^{\alpha_0}(dP_1)^{\alpha_1}\cdots(dP_k)^{\alpha_k}, \tag{*}$$

其中 $\alpha = (\alpha_0, \alpha_1, \cdots, \alpha_k)$ 属于单形

$$\Sigma_{k+1} = \left\{\alpha = (\alpha_0, \alpha_1, \cdots, \alpha_k):\ \alpha_i \geqslant 0,\ \sum_{i=0}^k \alpha_i = 1\right\}.$$

如同 $k=1$ 的情形, 要求对 (*) 中的积分赋予意义 (可利用强势测度的概念), 并且证明与引理 3 相类似的结论.

10. 设 $(\Sigma_k, \mathscr{B}(\Sigma_k))$ 是可测空间, 其中 Σ_k 是单形

$$\Sigma_k = \left\{ x = (x_1, \cdots, x_k):\ x_i \geqslant 0,\ \sum_{i=1}^k x_i = 1 \right\},$$

而 $\mathscr{B}(\Sigma_k)$ 是由该单形的子集所形成的博雷尔类.

设 $\mu = \mu(dx)$ 是 $(\Sigma_k, \mathscr{B}(\Sigma_k))$ 上的测度, 有 $\mu(\Sigma_k) < \infty$ 和

$$\int_{\Sigma_k} x_i \mu(dx) = 1, \quad i = 1, \cdots, k.$$

(在概率–统计试验中, 将具有这种性质的测度 μ 称为标准的.)

在数学分析中, 测度 μ 的海林格变换 $\mathbb{H}(\alpha;\ \mu)$ 如下式所定义: 对一切 $\alpha \in \Sigma_k$, 令

$$\mathbb{H}(\alpha;\ \mu) = \int_{\Sigma_k} x_1^{\alpha_1} \cdots x_k^{\alpha_k} \mu(dx).$$

试证明如下各命题:

(a) 如果 μ_1 与 μ_2 是两个标准测度, 使得对一切 $\alpha \in \Sigma_k$, 都有 $\mathbb{H}(\alpha;\ \mu_1) = \mathbb{H}(\alpha;\ \mu_2)$, 则有 $\mu_1 = \mu_2$.

(b) 标准测度序列 μ_n 弱收敛到标准测度 μ, 当且仅当对一切 $\alpha \in \Sigma_k$, 都有 $\mathbb{H}(\alpha;\ \mu_n) \to \mathbb{H}(\alpha;\ \mu)$, $n \to \infty$.

设 $\mathscr{E} = (\Omega, \mathscr{F};\ P_0, P_1, \cdots, P_k)$ 是一个概率–统计试验, Q 是一个可以控制测度 $P_0, P_1, \cdots, P_k$ 的强势测度, 而

$$f_i = \frac{dP_i}{dQ}, \quad i = 0, 1, \cdots, k.$$

在 $(\Sigma_k, \mathscr{B}(\Sigma_k))$ 上定义概率测度

$$\mu(A) = Q\{\omega:\ (f_0(\omega), \cdots, f_k(\omega)) \in A\}, \quad A \in \mathscr{B}(\Sigma_{k+1}).$$

证明, 测度 μ 是标准的, 并且

$$H(\alpha;\ \mathscr{E}) = \mathbb{H}(\alpha;\ \mu).$$

11. 设 $\mathscr{E} = (\Omega, \mathscr{F};\ P_0, P_1, \cdots, P_k)$ 是一个概率–统计试验, P_0 是可以控制测度 $P_1, \cdots, P_k$ 的强势测度. 令

$$z_i = \frac{dP_i}{dP_0}, \quad i = 1, \cdots, k.$$

在概率论中, 将按照下式所定义的 $\beta \in \Delta_k$ 的函数称为试验 $\mathscr{E}$ 的梅林 (Mellin) 变换:

$$M(\beta;\ \mathscr{E}) = \int_\Omega z_1^{\beta_1} \cdots z_k^{\beta_k} P_0(d\omega) \quad (= E_0(z_1^{\beta_1} \cdots z_k^{\beta_k})),$$

其中

$$\Delta_k = \left\{ \beta = (\beta_1, \cdots, \beta_k):\ 0 \leqslant \beta_i < 1,\ \sum_{i=1}^k \beta_i < 1 \right\}.$$

数学分析中, 对测度 ν 的梅林变换 $\mathbb{M}(\beta;\ \nu)$ 略有不同: 设 ν 是 $(\mathbb{R}_+^k, \mathscr{B}(\mathbb{R}_+^k))$ 上的一个概率测度, 其中

$$\mathbb{R}_+^k = \{x = (x_1, \cdots, x_k):\ x_i \geqslant 0,\ i = 1, \cdots, k\},$$

使得

$$\int_{\mathbb{R}_+^k} x_i \nu(dx) \leqslant 1.$$

($\mathbb{R}_+^k$ 上的这样的测度 ν 也称为标准的.) 此时, 根据定义, 令

$$\mathbb{M}(\beta;\ \nu) = \int_{\mathbb{R}_+^k} x_1^{\beta_1} \cdots x_k^{\beta_k} \nu(dx),$$

其中 $\beta = (\beta_1, \cdots, \beta_k) \in \Delta_k$.

试证明,

(a) 若 ν_1 与 ν_2 是两个标准概率测度, 使得对一切 $\beta \in \Delta_k$, 都有 $\mathbb{M}(\beta;\ \nu_1) = \mathbb{M}(\beta;\ \nu_2)$, 则有 $\nu_1 = \nu_2$.

(b) 标准测度 ν_n 的序列 (ν_n) 弱收敛到标准测度 ν, 当且仅当对一切 $\beta \in \Delta_k$, 都有 $\mathbb{M}(\beta;\ \nu_n) \to \mathbb{M}(\beta;\ \nu)$, $n \to \infty$.

(c) $M(\beta;\ \mathscr{E}) = \mathbb{M}(\beta;\ \nu)$.

12. 证明, 如果 $\alpha = (\alpha_0, \alpha_1, \cdots, \alpha_k) \in \Sigma_{k+1}$, 并且 $\alpha_0 > 0$, 则有

$$H(\alpha;\ \mathscr{E}) = M(\beta;\ \mathscr{E}),$$

其中 $\beta = (\beta_1, \cdots, \beta_k) = (\alpha_1, \cdots, \alpha_k)$.

再证明, 如果 $\alpha = (\alpha_0, \alpha_1, \cdots, \alpha_k) \in \Sigma_{k+1}$, 其中 $\alpha_0 > 0$, 并且 $L_i = \ln z_i$, $i = 1, \cdots, k$, 则有

$$H(\alpha;\ \mathscr{E}) = E_0 \exp\left\{ \sum_{i=1}^k \alpha_i L_i \right\},$$

亦即海林格变换 $H(\alpha;\ \mathscr{E})$ 重合于向量 $(L_1, \cdots, L_k)$ 关于测度 P_0 的拉普拉斯变换.

13. 设 $P = (p_{ab})$ 为随机矩阵 (见第一章第 12 节), $1 \leqslant a, b \leqslant N < \infty$. 量

$$D(P) = \frac{1}{2} \sup_{i,j} \sum_{k=1}^{N} |p_{ik} - p_{jk}|$$

称为矩阵 P 的多布鲁申 (Dobrushin) 遍历系数.

试证明,

(a) $D(P) = \sup_{i,j} \|p_{i\cdot} - p_{j\cdot}\|$($\|\cdot\|$ 为变差距离).

(b) $D(P) = 1 - \inf_{i,j} \sum_{k=1}^{N} (p_{ik} \wedge p_{jk})$ ①.

(c) 如果 P 与 Q 是两个随机矩阵, 则

$$D(PQ) \leqslant D(P)D(Q).$$

(d) 如果 $\mu = (\mu_1, \mu_2, \cdots, \mu_N)$ 与 $\nu = (\nu_1, \nu_2, \cdots, \nu_N)$ 为两个分布, 则

$$\|\mu P^n - \nu P^n\| \leqslant \|\mu - \nu\|(D(P))^n.$$

14. 设 P 与 Q 分别是随机变量 ξ 与 η 的分布, 则有如下的卡泊林格不等式(Kapling-inequality) 成立:

$$\mathbf{P}\{\xi = \eta\} \leqslant 1 - \frac{1}{2}\|P - Q\|.$$

(试比较第 8 题中的断言.) 特别地, 如果随机变量 ξ 与 η 分别具有密度函数 $p(x)$ 与 $q(x)$, 则

$$\mathbf{P}\{\xi = \eta\} \leqslant 1 - \frac{1}{2}\int_{\mathbb{R}} |p(x) - q(x)|dx.$$

试举例说明, 式中的等号可以成立.

15. 设 $X = (X_n)_{n \geqslant 0}$ 与 $Y = (Y_n)_{n \geqslant 0}$ 是定义在某个概率空间 $(\Omega, \mathscr{F}, \mathbf{P})$ 上的两个随机变量序列. 设 τ 是随机时, 使得对一切 $n \geqslant \tau(\omega)$, 都有 $X_n(\omega) = Y_n(\omega)$(随机时 τ 称为卡泊林格时). 证明如下的卡泊林格不等式成立:

$$\frac{1}{2}\|P_n - Q_n\| \leqslant \mathbf{P}\{\tau \geqslant n\},$$

其中 P_n 与 Q_n 分别是随机变量 X_n 与 Y_n 的概率分布.

16. 设 $f = f(x)$ 与 $g = g(x)$ 是 $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ 上的两个分布的密度函数. 证明,

(a) $\int |f(x) - g(x)|dx = 2\int (f(x) - g(x))^+ dx = 2\int (g(x) - f(x))^+ dx.$

(b) $\left(\int \sqrt{f(x)g(x)}dx\right)^2 \leqslant 2\int \min(f(x), g(x))dx.$

(c) $\int |f(x) - g(x)|dx \leqslant \sqrt{2K(f, g)}$, 其中 $K(f, g) = \int f(x) \ln \frac{f(x)}{g(x)}dx$ 是库尔贝克信息 (参阅第 3 题), 并假定具有密度 f 的分布 P_f 关于具有密度 g 的分布 P_g 绝对连续.

(d) $\int \min(f(x), g(x))dx \geqslant \frac{1}{2}e^{-K(f, g)}.$

17. 设随机向量 $X = (X_1, \cdots, X_k)$ 具有如下集合上的均匀分布:

$$T_k = \left\{x = (x_1, \cdots, x_k),\ x_i \geqslant 0,\ \sum_{i=1}^{k} x_i \leqslant 1\right\}.$$

证明, 向量 X 的概率分布密度 $f(x)$ 为

$$f(x) = k!, \quad x \in T_k.$$

18. 设 X 与 Y 为随机变量, 有 $\mathbf{E}X^2 < \infty$, $\mathbf{E}Y^2 < \infty$, 而 $\mathbf{cov}(X, Y) = \mathbf{E}(X - \mathbf{E}X)(Y - \mathbf{E}Y)$ 是它们的协方差. 分别以 $F(x, y)$, $F_1(x)$ 和 $F_2(y)$ 表示随机向量 (X, Y) 以及随机变量 X 与 Y 的分布函数. 试证明如下的霍夫丁公式:

$$\mathbf{cov}(X, Y) = \iint (F(x, y) - F_1(x)F_2(y))dxdy.$$

① 原书中的求和号上限为 ∞ —— 译者注.

§10. 概率测度的临近性和完全渐近可区分性

1. 设 $P^n = P_1^n \times \cdots \times P_n^n$, $\widetilde{P}^n = \widetilde{P}_1^n \times \cdots \times \widetilde{P}_n^n$, $n \geqslant 1$, 其中, P_k^n 与 $\widetilde{P}_k^n$ 分别为具有参数 $(a_k^n, 1)$ 和 $(\widetilde{a}_k^n, 1)$ 的高斯测度. 试求分别使得 $(\widetilde{P}^n) \lhd (P^n)$ 与 $(\widetilde{P}^n) \vartriangle (P^n)$ 成立的关于 (a_k^n) 和 $(\widetilde{a}_k^n)$ 的条件.

提示: 先通过直接计算, 证明

$$H(\alpha;\ P^n, \widetilde{P}^n) = \exp\left\{-\frac{\alpha(1-\alpha)}{2}\sum_{k=1}^{n}(\alpha_k^n - \widetilde{\alpha}_k^n)^2\right\},$$

然后利用 (11) 与 (12) 中的断言.

2. 设 $P^n = P_1^n \times \cdots \times P_n^n$, $\widetilde{P}^n = \widetilde{P}_1^n \times \cdots \times \widetilde{P}_n^n$, $n \geqslant 1$, 其中, P_k^n 与 $\widetilde{P}_k^n$ 为 $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ 上的概率测度, 有 $P_k^n(dx) = I_{[0,1]}(x)dx$, $\widetilde{P}_k^n(dx) = I_{[a_n, 1+a_n]}(x)dx$, $0 \leqslant a_n \leqslant 1$. 证明, $H(\alpha;\ P_k^n, \widetilde{P}_k^n) = 1 - a_n$ 和

$$(\widetilde{P}^n) \lhd (P^n) \quad \Leftrightarrow \quad (P^n) \lhd (\widetilde{P}^n) \quad \Leftrightarrow \quad \varlimsup_n n a_n = 0,$$

$$(\widetilde{P}^n) \vartriangle (P^n) \quad \Leftrightarrow \quad \varlimsup_n n a_n = \infty.$$

3. 设 $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n \geqslant 0})$ 是可测空间中的筛子, 亦即在 $(\Omega, \mathscr{F})$ 中赋予了一个 σ-代数流 $(\mathscr{F}_n)_{n \geqslant 0}$, 即有 $\mathscr{F}_0 \subseteq \mathscr{F}_1 \subseteq \cdots \subseteq \mathscr{F}$. 再设 $\mathscr{F} = \sigma\left(\bigcup_n \mathscr{F}_n\right)$. 假定 P 与 $\widetilde{P}$

是 $(\Omega, \mathscr{F})$ 中的两个概率测度, 并且 $P_n = P|\mathscr{F}_n$, $\widetilde{P}_n = \widetilde{P}|\mathscr{F}_n$, 即它们在 $\mathscr{F}_n$ 上的局限. 证明,

$$(\widetilde{P}_n) \lhd (P_n) \quad \Leftrightarrow \quad \widetilde{P} \ll P,$$
$$(\widetilde{P}_n) \lhd\rhd (P_n) \quad \Leftrightarrow \quad \widetilde{P} \sim P,$$
$$(\widetilde{P}_n) \bigtriangleup (P_n) \quad \Leftrightarrow \quad \widetilde{P} \perp P.$$

4. 设 $(\Omega, \mathscr{F}, \mathbf{P})$ 为概率空间, 其中 $\Omega = \{-1, 1\}^\infty$ 为二元制无限序列 $\omega = (\omega_1, \omega_2, \cdots)$ 的集合 (序列中的每一项只有 1 和 -1 两种情形); 对于任何 $a_i = \pm 1$, $i = 1, \cdots, n$, 则有 $\mathbf{P}\{\omega : (\omega_1, \cdots, \omega_n) = (a_1, \cdots, a_n)\} = 2^{-n}$. 再设 $\varepsilon_n(\omega) = \omega_n$, $n \geqslant 1$. (关于测度 $\mathbf{P}$, 序列 $\varepsilon = (\varepsilon_1, \varepsilon_2, \cdots)$ 是相互独立的伯努利随机变量序列, 有 $\mathbf{P}\{\varepsilon_n = 1\} = \mathbf{P}\{\varepsilon_n = -1\} = \frac{1}{2}$.)

构造序列 $S = (S_n)_{n\geqslant 0}$ 如下: $S_0 = 1$, $S_n = S_{n-1}(1 + \rho_n)$, 其中 $\rho_n = \mu_n + \sigma_n\varepsilon_n$, $\sigma_n > 0$, $\mu_n > \sigma_n - 1$(在这些条件下, 有 $S_n > 0$; 在金融数学中, S_n 表示时刻 n 时的资产总额; 参阅第七章第 11 节.)

令 $\mathbf{P}^n = \mathbf{P}|\mathscr{F}_n$, 其中 $\mathscr{F}_n = \sigma(\varepsilon_1, \cdots, \varepsilon_n)$. 在 $(\Omega, \mathscr{F})$ 中引入新的概率测度 $\widetilde{\mathbf{P}}$, 使得 $\varepsilon_1, \varepsilon_2, \cdots$ 仍然为相互独立的随机变量序列, 但是却有

$$\widetilde{\mathbf{P}}\{\varepsilon_n = 1\} = \frac{1}{2}(1 + b_n), \quad \widetilde{\mathbf{P}}\{\varepsilon_n = -1\} = \frac{1}{2}(1 - b_n),$$

其中 $b_n = -\mu_n/\sigma_n$.

(a) 证明, 序列 $S = (S_n)_{n\geqslant 0}$ 关于测度 $\widetilde{\mathbf{P}}$ 是鞅 (参阅第一章第 11 节和第七章第 1 节).

(b) 令 $\mathbf{P}^n = \mathbf{P}|\mathscr{F}_n$, 证明

$$H(\alpha;\ \widetilde{\mathbf{P}}^n,\ \mathbf{P}^n) = \prod_{k=1}^{n} \left[\frac{(1 + b_k)^\alpha + (1 - b_k)^\alpha}{2}\right].$$

并利用 (第 10 节中的) 定理 1, 由此导出

$$(\mathbf{P}^n) \lhd (\widetilde{\mathbf{P}}^n) \quad \Leftrightarrow \quad \sum_{k=1}^{\infty} b_k^2 < \infty.$$

(根据 "大" 金融市场理论, 前一结论表明, 条件 $\sum\limits_{k=1}^{\infty} \left(\frac{\mu_k}{\sigma_k}\right)^2 < \infty$ 是无渐近套汇的充分必要条件; 更详细的介绍见专著 [130] 第六章第 3 节.)

5. 与上一题不同, 现在令 $S_n = e^{h_1 + \cdots + h_n}$, $n \geqslant 1$, $S_0 = 1$, 其中 $h_k = \mu_k + \sigma_k\varepsilon_k$, $\sigma_k > 0$, $(\varepsilon_1, \varepsilon_2, \cdots)$ 为独立同分布的正态 $\mathcal{N}(0, 1)$ 随机变量序列.

令 $\mathscr{F}_n = \sigma(\varepsilon_1, \cdots, \varepsilon_n)$, $\mathbf{P}^n = \mathbf{P}|\mathscr{F}_n$, $n \geqslant 1$.

(a) 证明, 关于这样的测度 $\widetilde{\mathbf{P}}$, 其中 $\widetilde{\mathbf{P}}|\mathscr{F}_n = \widetilde{\mathbf{P}}^n$, 而 $d\widetilde{\mathbf{P}}^n = z_n d\mathbf{P}^n$, 并且

$$z_n = \exp\left\{-\sum_{k=1}^{n}\left(\frac{\mu_k}{\sigma_k} + \frac{\sigma_k}{2}\right)\varepsilon_k + \frac{1}{2}\sum_{k=1}^{n}\left(\frac{\mu_k}{\sigma_k} + \frac{\sigma_k}{2}\right)^2\right\},$$

序列 $(S_n)_{n\geqslant 0}$ 是一个鞅 (参阅第七章第 11 节).

(b) 再证明

$$H(\alpha;\ \widetilde{\mathbf{P}}^n, \mathbf{P}^n) = \exp\left\{-\frac{\alpha(1-\alpha)}{2}\sum_{k=1}^{n}\left(\frac{\mu_k}{\sigma_k} + \frac{\sigma_k}{2}\right)^2\right\},$$

与

$$(\mathbf{P}^n) \lhd (\widetilde{\mathbf{P}}^n) \quad \Leftrightarrow \quad \sum_{k=1}^{n}\left(\frac{\mu_k}{\sigma_k} + \frac{\sigma_k}{2}\right)^2 < \infty.$$

($\sum\limits_{k=1}^{n}\left(\frac{\mu_k}{\sigma_k} + \frac{\sigma_k}{2}\right)^2 < \infty$ 是在所考察的市场上无渐近套汇的条件; 见专著 [130] 第六章第 3 节 c.)

§11. 中心极限定理的收敛速度

1. 试证明不等式 (8).

2. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布的随机变量序列, 有 $\mathbf{E}\xi_1 = 0$, $\mathbf{D}\xi_1 = \sigma^2$ 和 $\mathbf{E}|\xi_1|^3 < \infty$. 现知, 在这些条件之下, 有如下的非一致估计: 对一切 $-\infty < x < \infty$, 有

$$|F_n(x) - \Phi(x)| \leqslant \frac{c\mathbf{E}|\xi_1|^3}{\sigma^3\sqrt{n}} \cdot \frac{1}{(1 + |x|)^3}.$$

试证明这一结果, 至少对伯努利随机变量给出证明. (关于这一命题及其证明, 以及关于下面的第 5 ~ 7 题的证明, 可以参阅专著 [88].)

3. 设 $(\xi_k)_{k\geqslant 1}$ 是独立同分布的随机变量序列, 均各以 1/2 的概率取值 ± 1. 令 $\varphi_2(t) = \mathbf{E}e^{it\xi_1} = \frac{1}{2}(e^{-it} + e^{it})$. 仿照拉普拉斯的做法, 证明 $(S_n = \xi_1 + \cdots + \xi_n)$:

$$\mathbf{P}\{S_{2n} = 0\} = \frac{1}{\pi}\int_0^\pi \varphi_2^n(t)dt \sim \frac{1}{\sqrt{\pi n}}, \quad n \to \infty.$$

4. 设 $(\xi_k)_{k\geqslant 1}$ 是独立同分布的随机变量序列, 均各以 $\frac{1}{2a+1}$ 的概率取值 $0, \pm 1, \cdots, \pm a$, 其中 $a \geqslant 1$. 令 $\varphi_{2a+1}(t) = \mathbf{E}e^{it\xi_1} = \frac{1}{2a+1}\left(1 + 2\sum\limits_{k=1}^{a}\cos tk\right)$.

如同上题, 仿照拉普拉斯的做法, 证明,

$$\mathbf{P}\{S_n = 0\} = \frac{1}{\pi}\int_0^\pi \varphi_{2a+1}^n(t)dt \sim \frac{\sqrt{3}}{\sqrt{2\pi(a+1)n}}, \quad n \to \infty.$$

特别地, 如果 $a=1$, 即随机变量 ξ_k 以相等的概率取值 $-1,0,1$ 时, 有

$$\mathbf{P}\{S_n=0\}\sim\frac{\sqrt{3}}{2\sqrt{\pi n}},\quad n\to\infty.$$

5. 证明, 如果 $F=F(x)$ 与 $G=G(x)$ 是两个分布函数, $f(t)$ 与 $g(t)$ 分别为它们的特征函数, 则有

$$\sup_x|F(x)-G(x)|\leqslant\frac{1}{4}\int_{-\pi}^{\pi}\left|\frac{f(t)-g(t)}{t}\right|dt.$$

6. 证明, 如果 $F=F(x)$ 与 $G=G(x)$ 是两个分布函数, $f(t)$ 与 $g(t)$ 分别为它们的特征函数, 而 $L(F,G)$ 是它们之间的莱维距离 (第 1 节第 4 题), 则对任何 $T\geqslant 2$, 都有

$$L(F,G)\leqslant\frac{1}{\pi}\int_0^T\left|\frac{f(t)-g(t)}{t}\right|dt+2e\frac{\ln T}{T}.$$

7. 设 $F_n(x)$ 是满足条件 $\mathbf{E}\xi_1=0$, $\mathbf{D}\xi_1=\sigma^2$ 和 $\mathbf{E}|\xi_1|^3=\beta_3<\infty$ 的独立同分布的随机变量序列的正则化部分和 $\frac{1}{\sigma\sqrt{\pi}}\sum_{j=1}^n\xi_j$ 的分布函数. 令 $\rho=\frac{\beta_3}{\sigma^3}$. 证明,

$$\varlimsup_n\inf_{(\widetilde{a},\widetilde{\sigma})}\sqrt{n}\left|F_n(x)-\Phi\left(\frac{x-\widetilde{a}}{\widetilde{\sigma}}\right)\right|\leqslant\frac{\rho}{\sqrt{2\pi}}.$$

§12. 泊松定理的收敛速度

1. 证明, 在 $\lambda_k=-\ln(1-p_k)$ 时, 有变差距离

$$\|B(p_k)-\Pi(\lambda_k)\|=2(1-e^{-\lambda_k}-\lambda_k e^{-\lambda_k})\leqslant\lambda_k^2,$$

因而, $\|B-\Pi\|\leqslant\sum_{k=1}^n\lambda_k^2$.

提示: 如果注意到

$$\|B(p_k)-\Pi(\lambda_k)\|=|(1-p_k)-e^{-\lambda_k}|+|p_k-\lambda_k e^{-\lambda_k}|$$
$$+e^{-\lambda_k}\sum_{i=2}^\infty\frac{\lambda_k^i}{i!}=2(1-e^{-\lambda_k}-\lambda_k e^{-\lambda_k}),$$

那么不等式 $\|B(p_k)-\Pi(\lambda_k)\|\leqslant\lambda_k^2$ 就可以由这一结果和周知的初等不等式推出: 对于 $x\geqslant 0$, 有 $2(1-e^{-x}-xe^{-x})\leqslant x^2$.

2. 试证明表达式 (9) 和 (10).

3. 设 $\xi_1,\cdots,\xi_n$ 为相互独立的伯努利随机变量, 有 $\mathbf{P}\{\xi_k=1\}=p_k$, $\mathbf{P}\{\xi_k=0\}=1-p_k$, $1\leqslant k\leqslant n$. 我们令 $\xi_0=0$, 并且对 $0\leqslant t\leqslant 1$, $\lambda>0$, 记

$$S_n(t)=\sum_{k=0}^{[nt]}\xi_k,$$

$$P_k^{(n)}(t)=\mathbf{P}\{S_n(t)=k\},\quad \pi_k(t)=\frac{(\lambda t)^k e^{-\lambda t}}{k!},\quad k=0,1,2,\cdots,$$

和

$$A_n(t)=\sum_{k=0}^{[nt]}p_k\quad(=\mathbf{E}S_n(t)).$$

证明, 对于概率 $P_k^{(n)}(t)$ 和 $\pi_k(t)$, 有如下各关系式成立:

$$\begin{aligned}P_0^{(n)}(t)&=1-\int_0^t P_0^{(n)}(s-)\,dA_n(s),\\ P_k^{(n)}(t)&=-\int_0^t[P_k^{(n)}(s-)-P_{k-1}^{(n)}(s-)]\,dA_n(s),\quad k\geqslant 1\end{aligned}\tag{$*$}$$

和

$$\begin{aligned}\pi_0(t)&=1-\int_0^t\pi_0(s-)\,d(\lambda s),\\ \pi_k(t)&=-\int_0^t[\pi_k(s-)-\pi_{k-1}(s-)]\,d(\lambda s),\quad k\geqslant 1.\end{aligned}\tag{$**$}$$

4. 试由上题中的 $(*)$ 和 $(**)$ 式, 推出如下关系式:

$$\begin{aligned}\sum_{k=0}^\infty|P_k^{(n)}(t)-\pi_k(t)|&\leqslant 2\int_0^t\sum_{k=0}^\infty|P_k^{(n)}(s-)-\pi_k(s-)|d(\lambda s)\\ &\quad+(2+4A_n(t))\max_{0\leqslant s\leqslant t}|A_n(s)-\lambda s|.\end{aligned}\tag{$***$}$$

5. 试利用第二章第 6 节第 51 题中的格朗沃尔–贝尔曼 (Gronwall-Bellman) 不等式, 从 $(***)$ 式推出 (参阅第 3 题中的记号)

$$\sum_{k=0}^\infty|P_k^{(n)}(t)-\pi_k(t)|\leqslant e^{2\lambda t}+(2+4A_n(t))\max_{0\leqslant s\leqslant t}|A_n(s)-\lambda s|.$$

并由此断言

$$\sum_{k=0}^\infty|\mathbf{P}\{S_n(1)=k\}-\pi_k(1)|$$
$$\leqslant\left(2+4\sum_{k=1}^n p_k\right)e^{2\lambda}\min_i\sup_{0\leqslant s\leqslant 1}\left|\sum_{k=0}^{[ns]}p_{i_k}-\lambda s\right|,$$

其中, min 系对正整数 $(1,\cdots,n)$ 的一切不同排列 $i=(i_1,\cdots,i_n)$ 所取, $p_{i_0}=0$.

再利用所得到的不等式, 证明, 如果 $\sum_{k=1}^n p_k=\lambda$, 则

$$\sum_{k=0}^\infty\left|\mathbf{P}\left\{\sum_{l=1}^n\xi_l=k\right\}-\frac{\lambda^k e^{-\lambda}}{k!}\right|$$
$$\leqslant C(\lambda)\min_i\sup_{0\leqslant s\leqslant 1}\left|\sum_{k=0}^{[ns]}p_{i_k}-\lambda s\right|\leqslant C(\lambda)\max_{1\leqslant k\leqslant n}p_k,$$

其中 $C(\lambda)=(2+4\lambda)e^{2\lambda}$.

§13. 数理统计的基本定理

1. 试证明 (18) 式.

2. 证明, (距离空间 $(D,\mathscr{D},\rho)$ 中的) 概率测度的弱收敛 $P^{(N)}\xrightarrow{w}P$ 蕴涵 $f(X^{(N)})\xrightarrow{d}f(X)$. (参阅《概率》第一卷第三章第 13 节中的符号.)

3. 试证明 (22) 式中的蕴涵关系.

4. 设 $\xi_1,\xi_2,\cdots$ 与 $\eta_1,\eta_2,\cdots$ 分别为具有连续的分布函数 $F=F(x)$ 和 $G=G(x)$ 的独立同分布的随机变量序列. 令

$$F_N(x;\,\omega)=\frac{1}{N}\sum_{k=1}^{N}I(\xi_k(\omega)\leqslant x),\quad G_N(x;\,\omega)=\frac{1}{N}\sum_{k=1}^{N}I(\eta_k(\omega)\leqslant x)$$

是经验分布函数.

我们令

$$D_{N,M}(\omega)=\sup_x|F_N(x;\,\omega)-G_M(x;\,\omega)|$$

和

$$D^+_{N,M}(\omega)=\sup_x(F_N(x;\,\omega)-G_M(x;\,\omega)).$$

人们在两样本场合下, 已知

$$\lim_{N,M\to\infty}\mathbf{P}\left\{\sqrt{\frac{NM}{N+M}}D_{N,M}(\omega)\leqslant y\right\}=K(y),\quad y>0,\tag{$*$}$$

其中 $K(y)$ 是柯尔莫戈洛夫分布 (参阅《概率》第一卷第三章第 13 节).

试根据对结论 (25) 的证明思路, 勾画证明结论 $(*)$ 以及断言 (27) 和 (28) 中的思路和基本步骤.

5. 考察"奥米咖平方" 统计量

$$\omega^2_N(\omega)=\int_{-\infty}^{\infty}|F_N(x;\,\omega)-F(x)|^2dF(x),\tag{$**$}$$

其中 $F=F(x)$ 是连续的分布函数. 证明, 与统计量 $D_N(\omega)$ 和 $D^+_N(\omega)$ 的情形一样, 统计量 $\omega^2_N(\omega)$ 的分布对于一切连续分布函数 $F=F(x)$ 都是相同的. 再证明

$$\mathbf{E}\omega^2_N(\omega)=\frac{1}{6N},\qquad \mathbf{D}\omega^2_N(\omega)=\frac{4N-3}{180N^3}.$$

6. 设 $\xi_1,\xi_2,\cdots$ 为独立同分布的随机变量序列, 有 $\mathbf{E}\xi_1=0$, $\mathbf{D}\xi_1=1$. 令

$$\mathscr{R}_n=\max_{k\leqslant n}\left(S_k-\frac{k}{n}S_n\right)-\min_{k\leqslant n}\left(S_k-\frac{k}{n}S_n\right).$$

证明

$$\frac{\mathscr{R}_n}{\sqrt{n}}\overset{d}{=}\max_{\mathbb{T}_n}|B_t-tB_1|-\min_{\mathbb{T}_n}|B_t-tB_1|\overset{d}{=}\max_{\mathbb{T}_n}B^\circ_t-\min_{\mathbb{T}_n}B^\circ_t,$$

其中, $\mathbb{T}_n=\{t=k/n;\ k=0,1,\cdots,n\}$, $B=(B_t)_{t\leqslant1}$ 是布朗运动, $B^\circ=(B^\circ_t)_{t\leqslant1}$ 是布朗桥, 而 $\overset{d}{=}$ 如同往常, 表示分布相同.

并且证明

$$\mathbf{E}\mathscr{R}_n\sim\sqrt{\frac{\pi}{2}n},\quad \mathbf{D}\mathscr{R}_n\sim\left(\frac{\pi^2}{6}-\frac{\pi}{2}\right)n.$$

(试比较第二章第 6 节第 87 题.)

7. 设 $F=F(x)$ 与 $G=G(x)$ 是两个分布函数, 而 $F^{-1}(t)=\inf\{x:\,F(x)>t\}$ 与 $G^{-1}(t)=\inf\{x:G(x)>t\}$.

记 $\mathfrak{F}_2=\left\{F:\ F\text{为分布函数, 且 }\int_{-\infty}^{\infty}x^2dF(x)<\infty\right\}$. 对于 $\mathfrak{F}_2$ 中的 F 与 G, 令

$$d_2(F,\,G)=\left(\int_0^1|F^{-1}(t)-G^{-1}(t)|^2dt\right)^{1/2}.$$

(a) 证明, $d_2=d_2(F,\,G)$ 是一个距离, 称为瓦塞尔斯坦 (Wasserstein) 距离, 并且 $(\mathfrak{F}_2,\,d_2)$ 是完备的距离空间.

(b) 证明, 如果 $\xi_1,\cdots,\xi_n$ 为独立同分布的随机变量, 具有属于 $\mathfrak{F}_2$ 的分布函数 F. 设 $\widehat{F}_n$ 是其经验分布函数, 则 ($\mathbf{P}$-a.s.地) 有

$$d_2(F,\,\widehat{F}_n)\to0,\quad n\to\infty.$$

(c) 证明, 距离 $d_2(F,\,G)$ 具有下述的卡泊林格性质(Kapling-property):

$$d_2(F,\,G)=\inf\mathbf{E}(\xi-\eta)^2,$$

其中 inf 系对所有的分别服从分布 F 与 G 的随机变量对 (ξ,η) 所取 $(F,\,G\in\mathfrak{F}_2)$.

8. 记 $\mathfrak{F}_1=\left\{F:\ F\text{为分布函数, 且 }\int_{-\infty}^{\infty}|x|dF(x)<\infty\right\}$. 对于 $\mathfrak{F}_1$ 中的 F 与 G, 令

$$d_1(F,\,G)=\int_0^1|F^{-1}(t)-G^{-1}(t)|dt.$$

(a) 证明, 这一被称为多布鲁申距离的度量具有如下性质:

$$d_1(F,\,G)=\int_{-\infty}^{\infty}|F(x)-G(x)|dx,$$

并且 $(\mathfrak{F}_1, d_1)$ 是完备的距离空间.

(b) 证明, 如果 $F, F_1, F_2, \cdots \in \mathfrak{F}_1$, 则 $d_1(F, F_n) \to 0,\ n \to \infty$, 当且仅当, $F_n \Rightarrow F$(淡收敛, 参阅第 1 节) 和 $\int |x| dF_n(x) \to \int |x| dF(x)$.

(c) 证明, 对于 $\mathfrak{F}_1$ 中的 F 与 G, 成立下述的卡治林格性质:

$$d_1(F, G) = \inf \mathbf{E}|\xi - \eta|,$$

其中 inf 系对所有的分别服从分布 F 与 G 的随机变量对 (ξ, η) 所取 $(F, G \in \mathfrak{F}_1)$.

(d) 证明, 如果 $\xi_1, \cdots, \xi_n$ 为独立同分布的随机变量, 具有属于 $\mathfrak{F}_1$ 的分布函数 F. 设 $\widehat{F}_n$ 是其经验分布函数, 则 ($\mathbf{P}$-a.s.地) 有

$$d_1(F, \widehat{F}_n) \to 0, \quad n \to \infty.$$

9. 设随机变量 X 具有分布函数 $F = F(x)$, $x \in \mathbb{R}$ 及其反函数 $F^{-1} = F^{-1}(u)$, $u \in [0, 1]$(参阅第 8 节第 8 题中的定义). 对一切 $0 < p < 1$, 称 $\kappa_p = F^{-1}(p)$ 为随机变量 X(或分布函数 $F = F(x)$) 的 p-分位数. (通常, 将 $F^{-1}(1/2)$ 称为中位数, 把 $F^{-1}(1/4)$ 与 $F^{-1}(3/4)$ 分别称为下、上 1/4 分位数.)

试给出使得 p-分位数 κ_p 是方程 $F(x) = p$ 的唯一根的条件.

10. 设 $X_1, \cdots, X_n$ 是具有分布函数 $F = F(x)$ 的独立同分布的随机变量. 以 $\widehat{F}_n = \widehat{F}_n(x)$ 表示根据 n 个变量 $X_1, \cdots, X_n$ 所构造的经验分布函数

$$\widehat{F}_n(x) = F_n(x; \omega) = \frac{1}{n} \sum_{k=1}^{n} I(X_k(\omega) \leqslant x)$$

(参阅 (1) 式).

证明, 如果 $\widehat{X}_1^{(n)}, \cdots, \widehat{X}_n^{(n)}$ 是由 n 个观察值 $X_1, \cdots, X_n$ 所得到的秩序统计量 (在第一章第 12 节第 8 题和第二章第 8 节第 19 题中, 将该统计量记为 $X_1^{(n)}, \cdots, X_n^{(n)}$). 则经验分布函数 $\widehat{F}_n = \widehat{F}_n(x)$ 具有下述表达式:

$$\widehat{F}_n(x) = \begin{cases} 0, & \text{如果 } x < \widehat{X}_1^{(n)}, \\ k/n, & \text{如果 } \widehat{X}_k^{(n)} \leqslant x < \widehat{X}_{k+1}^{(n)},\quad k = 1, \cdots, n-1,\ n > 1, \\ 1, & \text{如果 } x \geqslant \widehat{X}_n^{(n)}. \end{cases}$$

11. 设上一题中的条件成立, κ_p 是分布 $F = F(x)$ 的 p-分位数, 而 $\widehat{\kappa}_p(n) = \widehat{F}_n^{-1}(p)$ 是相应的经验分布 $\widehat{F}_n = \widehat{F}_n(x)$ 的 p-分位数. 证明, 如果 κ_p 是满足条件 $F(\kappa_p -) \leqslant p \leqslant F(\kappa_p)$ 的唯一值, 则当 $n \to \infty$ 时, 有

$$\widehat{\kappa}_p(n) \to \kappa_p \quad (\mathbf{P}\text{-a.s.}).$$

提示: 应当指出, 我们有 $\widehat{\kappa}_p(n) = \widehat{X}_{\lceil np \rceil}^{(n)}$, 并且证明, 对任何 $\delta > 0$, 都有

$$\mathbf{P}\left\{\underline{\lim}\, \widehat{X}_{\lceil np \rceil}^{(n)} > \kappa_p - \delta\right\} = \mathbf{P}\left\{\overline{\lim}\, \widehat{X}_{\lceil np \rceil}^{(n)} < \kappa_p + \delta\right\} = 1,$$

其中, $\lceil x \rceil$ 表示不超过实数 x 的最大整数.

12. 设 $X_1, X_2, \cdots$ 是独立同分布的随机变量序列, 其分布函数 $F = F(x)$ 连续. 对每个 $0 < p < 1$, 方程 $F(x) = p$ 都有唯一解 κ_p, 并且导函数 $F'(x)$ 在点 κ_p 处存在, 连续且为正. 以 $\widehat{X}_{\lceil np \rceil}^{(n)}$ 表示样本 p-分位数.

证明, 随机变量 $\sqrt{n}\left(\widehat{X}_{\lceil np \rceil}^{(n)} - \kappa_p\right)$ 依分布收敛到某个均值为 0, 方差为 $p(1-p)(F'(\kappa_p))^2$ 的高斯随机变量 N:

$$\sqrt{n}\left(\widehat{X}_{\lceil np \rceil}^{(n)} - \kappa_p\right) \xrightarrow{\text{law}} N.$$

提示: 设 $\xi_1, \cdots, \xi_n$ 为相互独立的在区间 $[0, 1]$ 上均匀分布的随机变量, 而 $\widehat{\xi}_1^{(n)}, \cdots, \widehat{\xi}_n^{(n)}$ 是相应的秩序统计量. 为证题中结论, 应当首先指出, 随机变量 $\widehat{X}_{\lceil np \rceil}^{(n)} - \kappa_p$ 与随机变量 $F^{-1}(\widehat{\xi}_{\lceil np \rceil}^{(n)}) - F^{-1}(p)$ 同分布, 然后利用引理 2, 并验证关于中心极限定理成立的林德伯格条件 (第 4 节定理 1).

第四章　独立随机变量之和与独立随机变量序列

§1. 0-1 律

1. 试证明定理 1 的推论.

提示: 利用随机变量 η 的分布函数仅可取 0 和 1 两个值的事实.

2. 证明, 如果 $(\xi_n)_{n\geqslant 1}$ 是独立随机变量序列, 则随机变量 $\overline{\lim}\,\xi_n$ 与 $\underline{\lim}\,\xi_n$ 都是 ($\mathbf{P}$-a.s.) 退化的.

提示: 证明, $\overline{\lim}\,\xi_n$ 与 $\underline{\lim}\,\xi_n$ 皆为 $\mathscr{X}$ 可测的.

3. 设 $(\xi_n)_{n\geqslant 1}$ 是独立随机变量序列, $S_n=\xi_1+\cdots+\xi_n$, 而常数 b_n 满足条件 $0<b_n\uparrow\infty$. 证明, 随机变量 $\overline{\lim}\,\frac{S_n}{b_n}$ 与 $\underline{\lim}\,\frac{S_n}{b_n}$ 都是 ($\mathbf{P}$-a.s.) 退化的.

提示: 在集合 $\{1,2,\cdots\}$ 中任意取定 N, 令

$$\widetilde{S}_n=\begin{cases}0, & n\leqslant N,\\ S_n-S_N, & n>N.\end{cases}$$

由性质 $\overline{\lim\limits_n}\,\frac{S_n}{b_n}=\overline{\lim\limits_n}\,\frac{\widetilde{S}_n}{b_n}$ 可知, 上极限 $\overline{\lim\limits_n}\,\frac{S_n}{b_n}$ 关于 $\bigcap\limits_{n=N}^{\infty}\mathscr{F}_n$ 可测, 再根据 N 的任意性, 即知其关于 $\mathscr{X}$ 可测. 由此不难得出所需的结论.

4. 设 $S_n=\xi_1+\cdots+\xi_n$, $n\geqslant 1$, 而 $\mathscr{X}(S)=\bigcap\mathscr{F}_n^{\infty}(S)$, 其中 $\mathscr{F}_n^{\infty}(S)=\sigma\{\omega: S_n,S_{n+1},\cdots\}$. 证明, 尾 σ-代数 $\mathscr{X}(S)$ 中的每个事件都是可交换的.

5. 设 $(\xi_n)_{n\geqslant 1}$ 为随机变量序列. 证明, 对任何实常数 c, 都有 $\{\overline{\lim}\,\xi_n\geqslant c\}\supseteq\overline{\lim}\,\{\xi_n\geqslant c\}$.

提示: 只需指出

$$\overline{\lim_n}\,\{\xi_n\geqslant c\}=\{\omega:\ \xi_n(\omega)\geqslant c\ \ \text{i.o.}\}.$$

6. 试举例说明, 存在尾事件 A (即属于 σ-代数 $\mathscr{X}=\bigcap\limits_{n=1}^{\infty}\mathscr{F}_n^{\infty}(\xi)$ 的事件, 其中 $\mathscr{F}_n^{\infty}(\xi)=\sigma(\xi_n,\xi_{n+1},\cdots)$, 而 $(\xi_n)_{n\geqslant 1}$ 为随机变量序列), 满足条件 $0<\mathbf{P}(A)<1$ (即其概率严格大于 0, 严格小于 1).

7. 设 $\xi_1,\xi_2,\cdots$ 为相互独立的随机变量, 有 $\mathbf{E}\xi_n=0$, $\mathbf{E}\xi_n^2=1$, $n\geqslant 1$, 并且对其有中心极限定理成立 (即有 $\mathbf{P}\{S_n/\sqrt{n}\leqslant x\}\to\Phi(x)$, $x\in\mathbb{R}$, 其中 $S_n=\xi_1+\cdots+\xi_n$). 证明, 此时

$$\overline{\lim_{n\to\infty}}\,n^{-1/2}S_n=+\infty\quad(\mathbf{P}\text{-a.s.}).$$

(特别地, 这一性质对满足条件 $\mathbf{E}\xi_1=0$, $\mathbf{E}\xi_1^2=1$ 的独立同分布的随机变量序列成立.)

8. 设 $\xi_1,\xi_2,\cdots$ 为独立同分布的随机变量, 有 $\mathbf{E}|\xi_1|>0$, 令 $S_n=\xi_1+\cdots+\xi_n$, $n\geqslant 1$. 证明,

$$\overline{\lim_{n\to\infty}}\,|S_n|=+\infty\quad(\mathbf{P}\text{-a.s.}).$$

9. 设 $\xi_1,\xi_2,\cdots$ 为独立同分布的随机变量, 有 $\mathbf{E}\xi_1=0$, $\mathbf{E}|\xi_1|>0$, 令 $S_n=\xi_1+\cdots+\xi_n$, $n\geqslant 1$. 证明, $\mathbf{P}$-a.s. 地有

$$\overline{\lim_{n\to\infty}}\,n^{-1/2}S_n=+\infty,\qquad \underline{\lim_{n\to\infty}}\,n^{-1/2}S_n=-\infty.$$

(试比较定理 2 中的断言; 亦可参阅第 7 题.)

10. 设 $\mathscr{F}_1,\mathscr{F}_2,\cdots$ 为相互独立的 σ-代数. 令 $\mathscr{G}=\bigcap\limits_{n=1}^{\infty}\sigma\left(\bigcup\limits_{j\geqslant n}\mathscr{F}_j\right)$. 证明, 对于每个集合 $G\in\mathscr{G}$, 都有 0-1 律成立 (即: 概率 $\mathbf{P}(G)$ 非 0 即 1).

11. 设 $A_1,A_2,\cdots$ 为相互独立的随机事件, 有 $\mathbf{P}(A_n)<1$, $n\geqslant 1$, $\mathbf{P}\left(\bigcup\limits_{n=1}^{\infty}A_n\right)=1$. 证明 $\mathbf{P}(\overline{\lim}\,A_n)=1$.

12. 设 $A_1,A_2,\cdots$ 为相互独立的随机事件, 记 $\mathbf{P}(A_n)=p_n$, $n\geqslant 1$. 根据 0-1 律可知, 概率 $\mathbf{P}(\overline{\lim}\,A_n)$ 与 $\mathbf{P}(\underline{\lim}\,A_n)$ 均非 0 即 1. 试通过 p_n, $n\geqslant 1$, 分别表述出使得 $\mathbf{P}(\underline{\lim}\,A_n)=0$, $\mathbf{P}(\underline{\lim}\,A_n)=1$, $\mathbf{P}(\overline{\lim}\,A_n)=0$ 和 $\mathbf{P}(\overline{\lim}\,A_n)=1$ 成立的条件.

13. 设 $\xi_1,\xi_2,\cdots$ 为非退化同分布的随机变量序列, 记 $S_n=\xi_1+\cdots+\xi_n$, $n\geqslant 1$. 证明,

(a) 对于每个博雷尔集合 $A\in\mathscr{B}(\mathbb{R})$, 都有 $\mathbf{P}\{S_n\in A\ \text{i.o.}\}=0$ 或 1.

(b) 仅存在 $\overline{\lim}\,S_n=+\infty$ ($\mathbf{P}$-a.s.) 或 $\overline{\lim}\,S_n=-\infty$ ($\mathbf{P}$-a.s.) 这两种可能性,

并且

$$\mathbf{P}\{\overline{\lim}\, S_n = +\infty\} = 1, \quad 如果 \quad \sum_{n=1}^{\infty} \frac{1}{n}\mathbf{P}\{S_n > 0\} = \infty,$$

$$\mathbf{P}\{\overline{\lim}\, S_n = -\infty\} = 1, \quad 如果 \quad \sum_{n=1}^{\infty} \frac{1}{n}\mathbf{P}\{S_n > 0\} < \infty.$$

(c) 如果 ξ_n 的分布对称, 则有 $\overline{\lim}\, S_n = +\infty$ 和 $\underline{\lim}\, S_n = -\infty$ ($\mathbf{P}$-a.s.)

14. 根据定理 1 的推论, 如果 $\xi_1, \xi_2, \cdots$ 为 (关于概率 $\mathbf{P}$) 相互独立的随机变量序列, 而 η 是关于由该序列所生成的尾 σ-代数 $\mathscr{X}$ 可测的随机变量, 则 η ($\mathbf{P}$-a.s. 地) 为常数, 即有 $\mathbf{P}\{\eta = C_{\mathbf{P}}\} = 1$, 其中 $C_{\mathbf{P}}$ 为某个常数. 现设 $\mathbf{Q}$ 是另一个概率测度, 关于该测度 $\xi_1, \xi_2, \cdots$ 仍为相互独立的随机变量序列, 那么当然也就有 $\mathbf{Q}\{\eta = C_{\mathbf{Q}}\} = 1$, 其中 $C_{\mathbf{Q}}$ 亦为某个常数. 试问, 是否必有 $C_{\mathbf{P}} = C_{\mathbf{Q}}$?

15. 设 $S_n = \xi_1 + \cdots + \xi_n$, $n \geqslant 1$, 其中 $\xi_1, \xi_2, \cdots$ 为相互独立的伯努利随机变量序列, 即有 $\mathbf{P}\{\xi_i = 1\} = \mathbf{P}\{\xi_i = -1\} = 1/2$, $i \geqslant 1$. 令 $\sigma_0 = \inf\{n \geqslant 1: \ S_n = 0\}$ (如果对一切 $n \geqslant 1$, 都有 $S_n \neq 0$, 则令 $\sigma_0 = \infty$). 记 $S_0 = 0$. 证明, 随机游动 $(S_n)_{n\geqslant 0}$ 在如下的意义下是自反的: 即有 $\mathbf{P}\{\sigma_0 < \infty\} = 1$. 并由此推出 $\mathbf{P}\{S_n = 0 \ \text{i.o.}\} = 1$.

提示: 利用第一章第 5 节第 7 题, 根据该题中的结论, 对任何 $n \geqslant 1$, 都有

$$\mathbf{P}\{S_1 \cdots S_{2n} \neq 0\} = 2^{-2n} C_{2n}^n.$$

16. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布随机变量序列, 有 $\mathbf{E}|\xi_1| < \infty$. 令 $S_n = \xi_1 + \cdots + \xi_n$, $n \geqslant 1$, 并设 $\mathbf{E}\xi_1 = 0$. 证明,

$$\underline{\lim}_{n\to\infty} |S_n| < \infty \quad (\mathbf{P}\text{-a.s.}).$$

(试比较第 8 题与第 9 题.)

17. 设 $X = (X_1, X_2, \cdots)$ 是可交换随机变量的无穷序列 (参阅第二章第 5 节第 4 题中的定义). 令 $\mathscr{X}_n = \sigma(X_n, X_{n+1}, \cdots)$, 而 $\mathscr{X} = \bigcap_n \mathscr{X}_n$ 是序列 X 所生成的尾 σ-代数.

证明, 对于任何有界的博雷尔函数 $g = g(x)$, 都有

$$\mathbf{E}[g(X_1) \,|\, \mathscr{X}] = \mathbf{E}[g(X_1) \,|\, \mathscr{X}_2] \quad (\mathbf{P}\text{-a.s.}).$$

再证明, 随机变量 $X_1, X_2, \cdots$ 关于尾 σ-代数 $\mathscr{X}$ 条件独立.

18. 设 $(X_1, \cdots, X_N)$ 是可交换随机变量的高斯向量. 证明, 存在服从标准正态分布的相互独立的随机变量 $\varepsilon_1, \cdots, \varepsilon_N$ ($\varepsilon_i \sim \mathcal{N}(0,1)$), 使得对一切 $1 \leqslant n \leqslant N$, 都有

$$X_n \overset{\text{law}}{=} a + b\varepsilon_n + c\sum_{i=1}^{N} \varepsilon_i,$$

其中, a, b 与 c 为某些常数.

19. 设 $(X_1, X_2, \cdots)$ 为可交换随机变量的无穷的高斯序列. 证明, 存在由独立同分布的随机变量构成的高斯序列 $(\varepsilon_0, \varepsilon_1, \cdots)$, 其中 $\varepsilon_i \sim \mathcal{N}(0,1)$, $i \geqslant 0$, 使得

$$X_n \overset{\text{law}}{=} a + b\varepsilon_0 + c\varepsilon_n, \quad n \geqslant 1.$$

20. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布的随机变量序列, 服从 $\mathbf{P}\{\xi_i > x\} = e^{-x}$, $x \geqslant 0$. 我们来考察事件 $A_n = \{\xi_n \geqslant h(n)\}$, $n \geqslant 1$, 其中 $h(n)$ 是如下各种函数中的任何一种: $c\ln n$, $\ln n + c\ln\ln n$ 或 $\ln n + \ln\ln n + c\ln\ln\ln n$.

证明,

$$\mathbf{P}\{A_n \ \text{i.o.}\} = \begin{cases} 0, & 如果\ c > 1, \\ 1, & 如果\ c \leqslant 1. \end{cases}$$

提示: 利用博雷尔–坎泰利引理.

21. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布的随机变量序列, 有 $\mathbf{P}\{\xi_n = 1\} = \mathbf{P}\{\xi_n = 0\} = 1/2$, $n \geqslant 1$. 我们来考察事件

$$A_n = \{\xi_{n+1} = 1, \ \cdots, \ \xi_{n+[\log_2 \log_2 n]} = 1\}, \quad n \geqslant 4.$$

(a) 证明, $\mathbf{P}\{A_n \ \text{i.o.}\} = 1$.

提示: 先对 $n = 2^m$, $m \geqslant 2$, 考察事件 A_n.

(b) 试求概率 $\mathbf{P}\{B_n \ \text{i.o.}\}$, 其中

$$B_n = \{\xi_{n+1} = 1, \ \cdots, \ \xi_{n+[\log_2 n]} = 1\}, \quad n \geqslant 2.$$

22. 设 $A_1, A_2, \cdots$ 为相互独立的事件. 我们来考察事件

$$B_{\leqslant x} = \left\{\omega: \ \lim_n \frac{1}{n}\sum_{k=1}^{n} I_{A_k} \leqslant x\right\}, \quad x \in \mathbb{R}.$$

证明, 对每个 $x \in \mathbb{R}$, 都有

$$\mathbf{P}\{B_{\leqslant x}\} = 0 \ \ 或 \ \ 1.$$

§2. 级数的收敛性

1. 设 $\xi_1, \xi_2, \cdots$ 为相互独立的随机变量序列. 试运用三级数定理, 证明,

(a) 如果 $\sum \xi_n^2 < \infty$ ($\mathbf{P}$-a.s.), 则 $\sum \xi_n$ 以概率 1 收敛, 当且仅当, 级数 $\sum \mathbf{E}\xi_n I(|\xi_n| \leqslant 1)$ 收敛.

(b) 如果级数 $\sum \xi_n$ ($\mathbf{P}$-a.s.) 收敛, 则 $\sum \xi_n^2 < \infty$ ($\mathbf{P}$-a.s.), 当且仅当

$$\sum \Big(\mathbf{E}|\xi_n| I(|\xi_n| \leqslant 1)\Big)^2 < \infty.$$

2. 设 $\xi_1, \xi_2, \cdots$ 为相互独立的随机变量序列. 证明, $\sum \xi_n^2 < \infty$ ($\mathbf{P}$-a.s.), 当且仅当,

$$\sum \mathbf{E}\frac{\xi_n^2}{1+\xi_n^2} < \infty.$$

提示: 利用三级数定理, 并注意

$$\sum \mathbf{E}\frac{\xi_n^2}{1+\xi_n^2} < \infty \quad \Leftrightarrow \quad \Big[\sum \mathbf{E}\xi_n^2 I(|\xi_n| \leqslant 1) < \infty \ \text{ 且 } \ \sum \mathbf{P}\{|\xi_n| > 1\} < \infty\Big].$$

3. 设 $\xi_1, \xi_2, \cdots$ 为相互独立的随机变量序列. 证明, 如下三个条件相互等价:

(i) 级数 $\sum \xi_n$ 以概率 1 收敛.

(ii) 级数 $\sum \xi_n$ 依概率收敛.

(iii) 级数 $\sum \xi_n$ 依分布收敛.

提示: 最方便的做法是: 依次证明 (i)$\Rightarrow$(iii)$\Rightarrow$(ii)$\Rightarrow$(i). 其中第一个蕴涵关系可由第二章第 10 节定理 2 推出; 蕴涵关系 (iii)$\Rightarrow$(ii) 需用反证法证明, 并要用到普罗霍洛夫定理; 为证蕴涵关系 (ii)$\Rightarrow$(i), 应当运用, 例如, 埃特麦迪不等式 (见第 22 题) 或者任何一个合适的不等式 (诸如, 斯科罗霍德不等式 (见第 4 节第 3 题), 奥坦韦安尼不等式 (见第七章第 3 节第 3 题)). 如果级数 $\sum \xi_n$ 依概率收敛, 则对任何 $\varepsilon > 0$, 存在 $m \in \mathbb{N} = \{1, 2, \cdots\}$, 使得只要 $n \geqslant m$, 就有

$$\max_{m \leqslant k \leqslant n} \mathbf{P}\{|S_n - S_k| > \varepsilon\} < \varepsilon.$$

利用上述不等式即可推出级数 $\sum \xi_n$ 以概率 1 收敛.

4. 试举例说明, 定理 1 和定理 2 中的一致有界性条件 (对某个 $c > 0$ 和一切 $n \geqslant 1$, 有 $\mathbf{P}\{|\xi_n| \leqslant c\} = 1$), 一般来说, 是不能去掉的.

提示: 考察独立随机变量序列 $\{\xi_n, n \geqslant 1\}$, 其中 ξ_n 的分布为

$$\mathbf{P}\{\xi_n = 0\} = 1 - \frac{2}{n^2}, \quad \mathbf{P}\{\xi_n = n\} = \mathbf{P}\{\xi_n = -n\} = \frac{1}{n^2}, \quad n \geqslant 1.$$

5. 设 $\xi_1, \cdots, \xi_n$ 为独立同分布随机变量, 有 $\mathbf{E}\xi_1 = 0$, $\mathbf{E}\xi_1^2 < \infty$. 令 $S_k = \xi_1 + \cdots + \xi_k$, $k \leqslant n$. 证明如下的与柯尔莫戈洛夫不等式单边情况相类似的马歇尔 (Marshall) 不等式

$$\mathbf{P}\left\{\max_{1 \leqslant k \leqslant n} S_k \geqslant \varepsilon\right\} \leqslant \frac{\mathbf{E}\,S_n^2}{\varepsilon^2 + \mathbf{E}\,S_n^2}, \quad \varepsilon \geqslant 0.$$

6. 设 $\xi_1, \xi_2, \cdots$ 为 (任意) 随机变量序列. 证明, 如果 $\sum\limits_{n \geqslant 1} \mathbf{E}|\xi_n| < \infty$, 则 $\sum\limits_{n \geqslant 1} \xi_n$ 以概率 1 绝对收敛.

7. 设 $\xi_1, \xi_2, \cdots$ 为相互独立的对称分布的随机变量. 证明,

$$\mathbf{E}\left[\left(\sum_{n \geqslant 1} \xi_n\right)^2 \wedge 1\right] \leqslant \sum_{n \geqslant 1} \mathbf{E}(\xi_n^2 \wedge 1).$$

8. 设 $\xi_1, \xi_2, \cdots$ 为相互独立的二阶矩有限的随机变量. 证明, 级数 $\sum \xi_n$ 在 L^2 中收敛, 当且仅当, 级数 $\sum \mathbf{E}\,\xi_n$ 与级数 $\sum \mathbf{D}\,\xi_n$ 都收敛.

9. 设 $\xi_1, \xi_2, \cdots$ 为相互独立的随机变量, 级数 $\sum \xi_n$ $\mathbf{P}$-a.s. 收敛. 证明, 该级数的值 $\mathbf{P}$-a.s. 地不依赖于求和顺序, 当且仅当, $\sum |\mathbf{E}\,(\xi_n;\ |\xi_n| \leqslant 1)| < \infty$.

10. 设 $\xi_1, \xi_2, \cdots$ 为相互独立的随机变量, 有 $\mathbf{E}\,\xi_n = 0$, $n \geqslant 1$ 及

$$\sum_n \mathbf{E}[\xi_n^2 I(|\xi_n| \leqslant 1) + |\xi_n| I(|\xi_n| > 1)] < \infty.$$

证明级数 $\sum\limits_n \xi_n$ ($\mathbf{P}$-a.s.) 收敛

11. 设 $A_1, A_2, \cdots$ 为相互独立的随机事件, 有 $\mathbf{P}(A_n) > 0$, $n \geqslant 1$, 且 $\sum\limits_n \mathbf{P}(A_n) = \infty$. 证明, 当 $n \to \infty$ 时, 有

$$\sum_{j=1}^n I(A_j) \Big/ \sum_{j=1}^n \mathbf{P}(A_j) \to 1 \quad (\mathbf{P}\text{-a.s.}).$$

12. 设 $\xi_1, \xi_2, \cdots$ 为相互独立的随机变量, 分别具有期望 $\mathbf{E}\,\xi_n$ 与方差 σ_n^2, 并且

$$\lim_{n \to \infty} \mathbf{E}\,\xi_n = c, \ \lim_{n \to \infty} \sigma_n^{-2} = \infty.$$

证明, 当 $n \to \infty$ 时, 有

$$\sum_{j=1}^n \frac{\xi_j}{\sigma_j^2} \Big/ \sum_{j=1}^n \frac{1}{\sigma_j^2} \to c \quad (\mathbf{P}\text{-a.s.}).$$

13. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布的标准指数分布随机变量 ($\mathbf{P}\{\xi_1 > x\} = e^{-\lambda}$, $x \geqslant 0$).

证明, 如果对于正数 a_n, $n \geqslant 1$, 有 $\sum\limits_{n \geqslant 1} a_n$ 收敛, 则级数 $\sum\limits_{n \geqslant 1} a_n \xi_n$ 不仅概率为 1 地收敛, 而且对一切 $p \geqslant 1$, 在 L^p 中收敛.

14. 设 $(T_1, T_2, \cdots)$ 为泊松过程的跳跃时刻序列 (参阅第七章第 10 节), 而 $\alpha \in (0, 1)$. 证明, 级数 $\sum\limits_{i=1}^\infty T_i^{-1/\alpha}$ 以概率 1 收敛.

15. 设 $(\xi_n)_{n \geqslant 1}$ 为独立随机变量序列, 其中 ξ_n 服从区间 $[-\frac{1}{n}, \frac{1}{n}]$ 上的均匀分布. 证明, ($\mathbf{P}$-a.s.地) 有:

(a) 级数 $\sum\limits_n \xi_n$ 概率为 1 地收敛.

(b) 级数 $\sum\limits_n |\xi_n| = \infty$.

提示: 利用二级数定理和三级数定理 (定理 2 和定理 3).

16. 定理 3 (三级数定理) 断言, 对于由相互独立的随机变量 $\xi_1, \xi_2, \cdots$ 所构成的级数 $\sum\limits_{n \geqslant 1} \xi_n$, 只要对某个 $c > 0$, 如下三个级数同时收敛, 那么就有级数 $\sum\limits_{n \geqslant 1} \xi_n$ ($\mathbf{P}$-a.s.)

收敛:

$$\sum_{n\geqslant 1}\mathbf{E}\,\xi_n^c,\qquad \sum_{n\geqslant 1}\mathbf{D}\,\xi_n^c,\qquad \sum_{n\geqslant 1}\mathbf{P}\{|\xi_n|>c\},$$

其中 $\xi_n^c=\xi_n I(|\xi_n|\leqslant c)$.

试举例说明, 上述三个级数 (对某个 $c>0$) 的收敛性中无论缺少哪一个, 都会破坏级数 $\sum_{n\geqslant 1}\xi_n$ 的 ($\mathbf{P}$-a.s.) 收敛性.

17. 设 $\xi_1,\xi_2,\cdots$ 为随机变量序列, 对某个 $r>0$, 有 $\sum_{k=1}^{\infty}\mathbf{E}\,|\xi_k|^r<\infty$. 证明, 当 $n\to\infty$ 时, 概率为 1 地有 $\xi_n\to 0$.

18. 设 $\xi_1,\xi_2,\cdots$ 为相互独立的对称伯努利随机变量, 即有 $\mathbf{P}\{\xi_k=1\}=\mathbf{P}\{\xi_k=-1\}=\frac{1}{2}$, $k\geqslant 1$. 证明, 级数 $\sum_{k=1}^{\infty}\frac{\xi_k}{2^k}$ 的概率分布为区间 $[-1,1]$ 上的均匀分布. (试比较第 10 节第 62 题和第二章第 12 节第 47 题.)

19. 设 $\xi_1,\xi_2,\cdots$ 为相互独立的对称分布随机变量. 证明, 如下三个条件相互等价:

(i) 级数 $\sum\xi_n$ 以概率 1 收敛.

(ii) 和 $\sum\xi_n^2<\infty$ ($\mathbf{P}$-a.s.).

(iii) 和 $\sum\mathbf{E}\,(\xi_n^2\wedge 1)<\infty$.

20. 设 ξ 为随机变量, 而 $\overline{\xi}$ 是它的对称化随机变量, 亦即 $\overline{\xi}=\xi-\widetilde{\xi}$, 其中 $\widetilde{\xi}$ 与 ξ 相互独立, 且与 ξ 同分布. (假设概率空间足够丰富.) 设 $\mu=\mu(\xi)$ 是随机变量 ξ 的按照意义 $\max(\mathbf{P}\{\xi>\mu\},\ \mathbf{P}\{\xi<\mu\})\leqslant\frac{1}{2}$ 所定义的中位数 (试比较第一章第 4 节第 23 题).

证明, 对一切 $a\geqslant 0$, 都有

$$\mathbf{P}\{|\xi-\mu|>a\}\leqslant 2\mathbf{P}\{|\overline{\xi}|>a\}\leqslant 4\mathbf{P}\left\{|\xi|>\frac{a}{2}\right\}.$$

21. 设 $\xi_1,\xi_2,\cdots$ 为独立随机变量序列, 有

$$\mathbf{P}\{\xi_n=1\}=2^{-n},\quad \mathbf{P}\{\xi_n=0\}=1-2^{-n}.$$

证明, 级数 $\sum_{n=1}^{\infty}\xi_n$ 以概率 1 收敛, 并且

$$\mathbf{P}\left\{\sum_{n=1}^{\infty}\xi_n=0\right\}=\prod_{n=1}^{\infty}(1-2^{-n})>0,$$

而

$$\mathbf{P}\left\{\sum_{n=1}^{\infty}\xi_n=1\right\}=\sum_{n=1}^{\infty}\frac{2^{-n}}{1-2^{-n}}\cdot\prod_{n=1}^{\infty}(1-2^{-n}).$$

22. 证明, 如果 $\xi_1,\xi_2,\cdots$ 是相互独立的随机变量序列, $S_m=\xi_1+\cdots+\xi_n$, $m\geqslant 1$, 则有如下的埃特麦迪不等式成立: 对任何 $\varepsilon>0$ 和 $n\geqslant 1$, 都有

$$\mathbf{P}\left\{\max_{1\leqslant m\leqslant n}|S_m|>4\varepsilon\right\}\leqslant 4\max_{1\leqslant m\leqslant n}\mathbf{P}\{|S_m|>\varepsilon\}.$$

23. 设 $\xi_1,\xi_2,\cdots$ 为相互独立的随机变量, 有 $\mathbf{E}\,\xi_k=0$, 并且对给定的 $h>0$, 有 $\mathbf{E}\,e^{h\xi_k}<\infty$, $k=1,\cdots,n$. 令 $S_k=\xi_1+\cdots+\xi_k$, $1\leqslant k\leqslant n$. 证明, 对一切 $\varepsilon>0$, 都有如下的指数型柯尔莫戈洛夫不等式成立:

$$\mathbf{P}\left\{\max_{1\leqslant k\leqslant n}S_k\geqslant\varepsilon\right\}\leqslant e^{-h\varepsilon}\mathbf{E}\,e^{hS_n}.$$

提示: 与证明柯尔莫戈洛夫不等式类似, 引入集合 $A=\left\{\max_{1\leqslant k\leqslant n}S_k\geqslant\varepsilon\right\}$, $A_k=\{S_i<\varepsilon,\ 1\leqslant i\leqslant k-1,\ S_k\geqslant\varepsilon\}$, $1\leqslant k\leqslant n$. 再由延森 (Jensen) 不等式得

$$\mathbf{E}\,e^{hS_n}\geqslant\mathbf{E}\,e^{hS_n}I_A=\sum_{k=1}^{n}\mathbf{E}\,e^{hS_n}I_{A_k}\geqslant\cdots\geqslant e^{h\varepsilon}\mathbf{P}(A).$$

24. 设 Y 为随机变量, $(Y_n)_{n\geqslant 1}$ 为随机变量序列, 当 $n\to\infty$ 时, 有 $Y_n\xrightarrow{\mathrm{d}}Y$ (依分布收敛). 又设 $\{N_t,\ t\geqslant 0\}$ 是与 $(Y_n)_{n\geqslant 1}$ 独立的正整数值随机变量族, 有 $N_t\xrightarrow{\mathbf{P}}\infty$, $t\to\infty$. 证明, $Y_{N_t}\xrightarrow{\mathrm{d}}Y$, $t\to\infty$.

提示: 以特征函数作为工具.

25. 设 Y 为随机变量, $(Y_n)_{n\geqslant 1}$ 为随机变量序列, 有

$$Y_n\to Y\quad(\mathbf{P}\text{-a.s.}),\quad n\to\infty,$$

$\{N_t,\ t\geqslant 0\}$ 是正整数值随机变量族 (与第 24 题不同, 此处不假定 $\{N_t,\ t\geqslant 0\}$ 与 $(Y_n)_{n\geqslant 1}$ 独立).

证明如下各命题:

(a) 如果 $N_t\to\infty$ ($\mathbf{P}$-a.s.), 则 $Y_{N_t}\to Y$ ($\mathbf{P}$-a.s.), $t\to\infty$.

(b) 如果 $N_t\to N$ ($\mathbf{P}$-a.s.), 则 $Y_{N_t}\to Y_N$ ($\mathbf{P}$-a.s.), $t\to\infty$.

(c) 如果 $N_t\xrightarrow{\mathbf{P}}\infty$, 则 $Y_{N_t}\xrightarrow{\mathbf{P}}Y$, $t\to\infty$.

提示: 在证明命题 (c) 时需要利用如下事实: 在依概率收敛的序列中存在几乎必然收敛的子列.

26. 设 $\xi_1,\xi_2,\cdots$ 是相互独立的伯努利随机变量序列, 有 $\mathbf{P}\{\xi_n=\pm 1\}=1/2$, $n\geqslant 1$. 证明, 随机变量 $X=\sum_{n=1}^{\infty}\frac{\xi_n}{n}$ 是确定的, 并且其分布函数具有密度.

27. 设 $\xi_1,\xi_2,\cdots$ 是相互独立的伯努利随机变量序列, 有 $\mathbf{P}\{\xi_n=0\}=\mathbf{P}\{\xi_n=1\}=1/2$, $n\geqslant 1$. 又设 $a_n>0$, $b_n>0$, $a_n+b_n=1$, $n\geqslant 1$, 并且

$$X_n=2a_n^{\xi_n}b_n^{1-\xi_n}.$$

证明, 如下各命题相互等价:

(i) $\prod_{n=1}^{\infty} X_n$ 几乎必然收敛 (亦即 $\lim_N \prod_{n=1}^{N} X_n$ 几乎必然存在且不为零).

(ii) $\prod_{n=1}^{\infty}(2-X_n)$ 几乎必然收敛.

(iii) $\prod_{n=1}^{\infty} a_n b_n$ 收敛.

提示: 在证明 (iii)⇒(i) 时, 应当分析 $\mathbf{E}\ln X_n$ 与 $\mathbf{D}\ln X_n$, 并利用三级数定理.

28. 设 $\xi_1,\xi_2,\cdots$ 为独立同分布的随机变量, 具有 (柯西) 密度函数 $f(x)=\frac{1}{\pi(1+x^2)}$, $x\in\mathbb{R}$. 证明, 对于任何常数 m, 都不能成立 $\frac{1}{n}\sum_{i=1}^{n}\xi_i \xrightarrow{\mathbf{P}} m$.

29. $\xi_1,\xi_2,\cdots$ 为独立同分布的随机变量, 有 $\mathbf{E}\xi_i=\mu$ 和 $\mathbf{D}\xi_i<\infty$. 证明, 当 $n\to\infty$ 时, 有

$$\frac{1}{C_n^2}\sum_{1\leqslant i<j\leqslant n}\xi_i\xi_j \xrightarrow{\mathbf{P}} \mu,$$

其中, C_n^2 为由 n 个不同元素中取出 2 个的所有不同的组合数目 $(=n(n-1)/2)$.

30. 设 $\xi_1,\xi_2,\cdots$ 是相互独立的伯努利随机变量序列, 有 $\mathbf{P}\{\xi_n=0\}=\mathbf{P}\{\xi_n=1\}=1/2$, $n\geqslant 1$. 对每个 $n\geqslant 1$, 以 Z_n 表示在序列 $\xi_1,\cdots,\xi_n$ 中所出现的连续的 1 所形成的片断的最大长度. 证明, 概率为 1 地有

$$\lim_n \frac{Z_n}{\ln n}=1.$$

提示: 分别证明, 概率为 1 地有 $\varliminf_n \frac{Z_n}{\ln n}\geqslant 1$ 和 $\varlimsup_n \frac{Z_n}{\ln n}\leqslant 1$.

31. 设 $\xi_1,\xi_2,\cdots$ 为相互独立的伯努利随机变量, 有 $\mathbf{P}\{\xi_n=1\}=p_n$, $\mathbf{P}\{\xi_n=0\}=1-p_n$, $n\geqslant 1$.

(a) 证明, 如果 $\sum_{k=1}^{\infty} p_k p_{k+1}<\infty$, 则级数 $\sum_{k=1}^{\infty}\xi_k\xi_{k+1}$ 以概率 1 收敛.

(b) 令 $p_n=1/n$, $n\geqslant 1$. 证明如下的狄厄康尼斯的结论: 随机变量 $S=\sum_{n=1}^{\infty}\xi_n\xi_{n+1}$ 具有参数 $\lambda=1$ 的泊松分布.

§3. 强大数定律

1. 证明, $\mathbf{E}\xi^2<\infty$, 当且仅当, $\sum_{n=1}^{\infty} n\mathbf{P}\{|\xi|>n\}<\infty$.

提示: 证明

$$\sum_{n=1}^{\infty} n\mathbf{P}\{|\xi|>n\}\leqslant \mathbf{E}\xi^2\leqslant 1+4\sum_{n=1}^{\infty} n\mathbf{P}\{|\xi|>n\}.$$

2. 设 $\xi_1,\xi_2,\cdots$ 为独立同分布的随机变量序列. 证明马钦凯维奇–济格蒙德 (Marcinkiewicz-Zigmund) 强大数定律 (简称强大数律) 的正确性: 如果对某个 $0<\alpha<1$, 有 $\mathbf{E}|\xi_1|^\alpha<\infty$, 则有 $\frac{S_n}{n^{1/\alpha}}\to 0$ ($\mathbf{P}$-a.s.); 而如果对某个 $1\leqslant\beta<2$, 有 $\mathbf{E}|\xi_1|^\beta<\infty$, 则有 $\frac{S_n-n\mathbf{E}\xi_1}{n^{1/\beta}}\to 0$ ($\mathbf{P}$-a.s.).

3. 设 $\xi_1,\xi_2,\cdots$ 为独立同分布的随机变量序列, 有 $\mathbf{E}|\xi_1|=\infty$. 证明, 对任何数列 $(a_n)_{n\geqslant 1}$, 都有

$$\varlimsup_n \left|\frac{S_n}{n}-a_n\right|=\infty \quad (\mathbf{P}\text{-a.s.}).$$

4. 区间 $[0,1)$ 中的所有有理数 (在第 4 小节例 2 的意义下) 是否都是正则的?

5. 考察区间 $[0,1)$ 中的数的十进制展开式 $\omega=0.\omega_1\omega_2\cdots$.

(a) 将第 4 小节中关于二进制展开式所建立的强大数律改述为适合于现在场合的形式.

(b) 在十进制展开式中, 有理数是否都是正则的 (意即当 $n\to\infty$ 时, 对每个 $i=0,1,\cdots,9$, 都有 $\frac{1}{n}\sum_{k=1}^{n} I(\xi_k(\omega)=i)\to\frac{1}{10}$) ($\mathbf{P}$-a.s.)?

(c) 证明 (由恰姆别尔诺伊所提出的命题): 数

$$\omega=0.123456789101112\cdots$$

(在十进制展开式意义下, 参阅例 2) 是正则的, 其中 ω 中的小数部分由所有的十进制正整数依次排列而成.

6. (埃特麦迪) 证明, 如果将 $\xi_1,\xi_2,\cdots$ 相互独立的条件换为两两独立, 定理 3 中的结论仍然成立.

7. 证明, 在定理 3 的条件下, 亦可成立依平均收敛性 ($\mathbf{E}\left|\frac{S_n}{n}-m\right|\to 0$, $n\to\infty$).

8. 设 $\xi_1,\xi_2,\cdots$ 为独立同分布的随机变量, 有 $\mathbf{E}|\xi_1|^2<\infty$. 证明

$$n\mathbf{P}\{|\xi_1|\geqslant\varepsilon\sqrt{n}\}\to 0 \quad 和 \quad \frac{1}{\sqrt{n}}\max_{k\leqslant n}|\xi_k|\xrightarrow{\mathbf{P}} 0.$$

(试比较第二章第 10 节第 41 题.)

9. 试举例说明, 存在这样的独立随机变量序列 $\xi_1,\xi_2,\cdots$, 对其依概率存在极限 $\lim_{n\to\infty}\frac{S_n}{n}$, 但是该极限却非以概率 1 存在.

提示: 考察独立随机变量序列 $\xi_1,\xi_2,\cdots$, 其中 ξ_n 的分布为

$$\mathbf{P}\{\xi_n=0\}=1-\frac{1}{n\ln n},\quad \mathbf{P}\{\xi_n=\pm n\}=\frac{1}{2n\ln n}.$$

对其有 $\mathbf{E}S_n^2\leqslant\frac{n^2}{\ln n}$ 和 $\sum_{n=1}^{\infty}\mathbf{P}\{|\xi_n|\geqslant n\}=1$, 因而由博雷尔–坎泰利引理知 $\mathbf{P}\{|\xi_n|\geqslant n \text{ i.o.}\}=1$.

10. 设 $\xi_1, \xi_2, \cdots$ 为独立随机变量序列, 有 $\mathbf{P}\{\xi_n = \pm n^a\} = 1/2$. 证明, 对该序列有强大数律成立, 当且仅当, $a < 1/2$.

11. 证明, 可以将柯尔莫戈洛夫强大数律表述为: 如果 $\xi_1, \xi_2, \cdots$ 是独立同分布随机变量序列, 则

$$\mathbf{E}|\xi_1| < \infty \quad \Leftrightarrow \quad n^{-1}S_n \to \mathbf{E}\,\xi_1 \quad (\mathbf{P}\text{-a.s.}),$$
$$\mathbf{E}|\xi_1| = \infty \quad \Leftrightarrow \quad \overline{\lim_n} n^{-1}S_n = \infty \quad (\mathbf{P}\text{-a.s.}).$$

证明, 如果将独立性假设换为两两独立, 则上述第一个命题仍然成立.

12. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布随机变量序列. 证明,

$$\mathbf{E}\sup_n \left|\frac{\xi_n}{n}\right| < \infty \quad \Leftrightarrow \quad \mathbf{E}\,|\xi_1| \ln^+ |\xi_1| < \infty.$$

13. 令 $S_n = \xi_1 + \cdots + \xi_n,\ n \geqslant 1$, 其中 $\xi_1, \xi_2, \cdots$ 是独立同分布的随机变量序列. 证明, 对任何 $\alpha \in (0, 1/2]$, 都有如下三个命题之一成立:

(a) $n^{-\alpha}S_n \to \infty$ ($\mathbf{P}$-a.s.).

(b) $n^{-\alpha}S_n \to -\infty$ ($\mathbf{P}$-a.s.).

(c) $\overline{\lim_n}\ n^{-\alpha}S_n = \infty,\ \underline{\lim_n}\ n^{-\alpha}S_n = -\infty$ ($\mathbf{P}$-a.s.).

14. 令 $S_n = \xi_1 + \cdots + \xi_n,\ n \geqslant 1,\ S_0 = 0$, 而 $\xi_1, \xi_2, \cdots$ 为独立同分布的随机变量序列. 证明,

(a) 对任何 $\varepsilon > 0$, 都有

$$\sum_{n=1}^{\infty} \mathbf{P}\{|S_n| \geqslant n\varepsilon\} < \infty \quad \Leftrightarrow \quad \mathbf{E}\,\xi_1 = 0, \quad \mathbf{E}\,\xi_1^2 < \infty.$$

(b) 如果 $\mathbf{E}\xi_1 < 0$, 则对 $p > 1$, 有

$$\mathbf{E}\left(\sup_{n\geqslant 0} S_n\right)^{p-1} < \infty \quad \Leftrightarrow \quad \mathbf{E}\,(\xi_1^+)^p < \infty.$$

(c) 如果 $\mathbf{E}\,\xi_1 = 0$, 而 $1 < p \leqslant 2$, 则对某个常数 C_p, 有

$$\sum_{n=1}^{\infty} \mathbf{P}\left\{\max_{k\leqslant n} S_k \geqslant n\right\} \leqslant C_p \mathbf{E}\,|\xi_1|^p, \quad \sum_{n=1}^{\infty} \mathbf{P}\left\{\max_{k\leqslant n} |S_k| \geqslant n\right\} \leqslant 2C_p \mathbf{E}\,|\xi_1|^p.$$

(d) 如果 $\mathbf{E}\,\xi_1 = 0,\ 0 < \mathbf{E}\,\xi_1^2 = \sigma^2 < \infty$ ①, 而 $M(\varepsilon) = \sup_{n\geqslant 0}(S_n - n\varepsilon),\ \varepsilon > 0$, 则有

$$\lim_{n\to\infty} \varepsilon M(\varepsilon) = \frac{\sigma^2}{2}.$$

①原文为: $\mathbf{E}\,\xi_1 = 0,\ \mathbf{E}\,\xi_1^2 < \infty, \cdots\cdots$ —— 译者注.

15. (补充定理 2) 设 $\xi_1, \xi_2, \cdots$ 为独立随机变量序列, 有

$$\mathbf{P}\{\xi_n = 1\} = \mathbf{P}\{\xi_n = -1\} = \frac{1}{2}(1 - 2^{-n}),$$
$$\mathbf{P}\{\xi_n = 2^n\} = \mathbf{P}\{\xi_n = -2^n\} = 2^{-(n+1)}.$$

证明, 尽管此处有 $\sum\limits_{n=1}^{\infty} \frac{\mathbf{D}\xi_n}{n^2} = \infty$ (试比较定理 2 中的 (3) 式), 却 ($\mathbf{P}$-a.s.地) 有

$$\frac{\xi_1 + \cdots + \xi_n}{n} \to 0,$$

亦即仍有强大数律成立 (换言之, 仍有 (4) 式成立. 我们指出, 在现在的情况下, 有 $\mathbf{E}\,\xi_n = 0,\ n \geqslant 1$).

16. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布的随机变量序列, 有 $\mathbf{E}\,|\xi_1| = \infty$. 证明, 此种情况下, 至少成立如下一种性质:

$$\mathbf{P}\left\{\overline{\lim_n}\frac{1}{n}\sum_{k=1}^{n}\xi_k = +\infty\right\} = 1 \quad \text{或} \quad \mathbf{P}\left\{\underline{\lim_n}\frac{1}{n}\sum_{k=1}^{n}\xi_k = -\infty\right\} = 1.$$

17. 作为柯尔莫戈洛夫强大数律的推广, 证明如下的洛埃韦 (Loéve) 的结果: 如果 $\xi_1, \xi_2, \cdots$ 为独立随机变量序列, 满足条件

$$\sum_{n=1}^{\infty} \frac{\mathbf{E}\,|\xi_n|^{\alpha_n}}{n^{\alpha_n}} < \infty,$$

其中 $0 < \alpha_n \leqslant 2$, 并且当 $1 \leqslant \alpha_n \leqslant 2$ 时, 有 $\mathbf{E}\,\xi_n = 0$, 则概率为 1 地有 $\frac{1}{n}\sum\limits_{i=1}^{n}\xi_i \to 0$.

18. 试举例说明, 存在这样的独立随机变量序列 $\xi_1, \xi_2, \cdots$, 满足条件 $\mathbf{E}\,\xi_n = 0,\ n \geqslant 1$, 但是却有

$$\frac{1}{n}\sum_{i=1}^{n}\xi_i \to -\infty \quad (\mathbf{P}\text{-a.s.}).$$

提示: 考察这样的随机变量 ξ_n, 其分布为

$$\mathbf{P}\{\xi_n = -n\} = 1 - n^{-2}, \quad \mathbf{P}\{\xi_n = n^3 - n\} = n^{-2}, \quad n \geqslant 1.$$

19. 设 $\xi_1, \xi_2, \cdots$ 为独立随机变量序列, 有 $\mathbf{E}\,\xi_k = 0,\ k \geqslant 1$. 令

$$\xi_k^{(n)} = \begin{cases} \xi_k, & \text{如果 } |\xi_k| \leqslant n, \\ 0, & \text{如果 } |\xi_k| > n. \end{cases}$$

证明, 如下的关于大数律成立的 (柯尔莫戈洛夫) 条件: 为了

$$\frac{1}{n}\sum_{k=1}^{n}\xi_k \xrightarrow{\mathbf{P}} 0,$$

必须且只需, 当 $n \to \infty$ 时, 有

$$\sum_{k=1}^{n} \mathbf{P}\{|\xi_k| > n\} \to 0,$$
$$\frac{1}{n}\sum_{k=1}^{n} \mathbf{E}\,\xi_k^{(n)} \to 0,$$
$$\frac{1}{n^2}\sum_{k=1}^{n} \mathbf{D}\,\xi_k^{(n)} \to 0.$$

试举例说明, 上述最后一个条件 (尤其在涉及必要性时), 不能换为条件:

$$\frac{1}{n^2}\sum_{k=1}^{n} \mathbf{E}\,(\xi_k^{(n)})^2 \to 0.$$

20. 设 $N = (N_t)_{t\geqslant 0}$ 为更新过程 (参阅第 4 小节例 4): $N_t = \sum_{n=1}^{\infty} I(T_n \leqslant t)$, 其中 $T_n = \sigma_1 + \cdots + \sigma_n$, 而 $(\sigma_n)_{n\geqslant 1}$ 为独立同分布的随机变量序列, 有 $\mathbf{E}\,\sigma_1 = \mu$, $0 < \mu < \infty$. 根据强大数律, 有 $\frac{N_t}{t} \to \frac{1}{\mu}$ ($\mathbf{P}$-a.s.). 证明, 对任何 $r > 0$, 都有

$$\mathbf{E}\left(\frac{N_t}{t}\right)^r \to \frac{1}{\mu^r}.$$

(这一结果在 $\mu = \infty$ 时仍然成立, 此时 $1/\mu = 0$.)

21. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布的随机变量序列, $S_n = \xi_1 + \cdots + \xi_n$, 而 $\{N_t,\ t \geqslant 0\}$ 是取值于集合 $\{1, 2, \cdots\}$ 的随机变量族, 有 $N_t \to \infty$ ($\mathbf{P}$-a.s.), $t \to \infty$.

证明,

(a) 如果 $\mathbf{E}\,|\xi_1|^r < \infty$, $r > 0$, 则

$$\frac{\xi_{N_t}}{(N_t)^{1/r}} \to 0 \quad (\mathbf{P}\text{-a.s.}), \quad t \to \infty,$$

如果还有 $N_t/t \to \lambda$ ($\mathbf{P}$-a.s.), 其中 $0 < \lambda < \infty$, 则

$$\frac{\xi_{N_t}}{t^{1/r}} \to 0 \quad (\mathbf{P}\text{-a.s.}), \quad t \to \infty.$$

(b) 如果 $\mathbf{E}\,|\xi_1|^r < \infty$, $0 < r < 2$, 并且当 $1 \leqslant r < 2$ 时, 有 $\mathbf{E}\,\xi_1 = 0$, 则

$$\frac{S_{N_t}}{(N_t)^{1/r}} \to 0 \quad (\mathbf{P}\text{-a.s.}), \quad t \to \infty,$$

如果还有 $N_t/t \to \lambda$ ($\mathbf{P}$-a.s.), 其中 $0 < \lambda < \infty$, 则

$$\frac{S_{N_t}}{t^{1/r}} \to 0 \quad (\mathbf{P}\text{-a.s.}), \quad t \to \infty.$$

(c) 如果 $\mathbf{E}\,|\xi_1| < \infty$, 且 $\mathbf{E}\,\xi_1 = \mu$, 则

$$\frac{S_{N_t}}{N_t} \to \mu \quad (\mathbf{P}\text{-a.s.}), \quad t \to \infty,$$

如果还有 $N_t/t \to \lambda$ ($\mathbf{P}$-a.s.), 其中 $0 < \lambda < \infty$, 则

$$\frac{S_{N_t}}{t} \to \mu\lambda \quad (\mathbf{P}\text{-a.s.}), \quad t \to \infty.$$

提示: 为证 (a), 可以利用博雷尔–坎泰利引理和第 2 节第 25 题中的命题 (a). 为证 (b), 需利用马钦凯维奇–济格蒙德强大数律 (第 2 题). 为证 (c), 应当利用柯尔莫戈洛夫强大数律 (定理 3) 和第 2 节第 25 题中的命题 (a).

22. 设 $f = f(x)$ 为区间 $(0, \infty)$ 上的有界连续函数. 证明, 对一切 $a > 0$ 和任何 $x > 0$, 都有

$$\lim_{n\to\infty}\sum_{k=1}^{\infty} f\left(x + \frac{k}{n}\right) e^{-an}\frac{(an)^k}{k!} = f(x+a).$$

23. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布的随机变量, 有 $\mathbf{E}\,|\xi_1| < \infty$, $\mathbf{E}\,\xi_1 = \mu$. 证明, 当 $n \to \infty$ 时, 有

(a) $\dfrac{\ln n}{n}\sum_{k=2}^{n}\dfrac{\xi_k}{\ln k} \to \mu$ ($\mathbf{P}$-a.s.).

(b) $n^{\alpha-1}\sum_{k=1}^{n}\dfrac{\xi_k}{k^{\alpha}} \to \mu$ ($\mathbf{P}$-a.s.), 其中 $0 < \alpha < 1$.

§4. 重对数定律

1. 设 $\xi_1, \xi_2, \cdots$ 为独立随机变量序列, $\xi_n \sim \mathcal{N}(0, 1)$. 证明,

(a) $\mathbf{P}\left\{\overline{\lim_n}\dfrac{\xi_n}{\sqrt{2\ln n}} = 1\right\} = 1.$

(b) $\mathbf{P}\{\xi_n > a_n \ \text{ i.o.}\} = \begin{cases} 0, & \text{如果 } \sum_n \mathbf{P}\{\xi_1 > a_n\} < \infty, \\ 1, & \text{如果 } \sum_n \mathbf{P}\{\xi_1 > a_n\} = \infty. \end{cases}$

提示: (a) 固定 $c > 0$, 证明, 根据 (10) 式 (《概率》第二卷 p24), 对于事件 $A_n = \{\xi_n > c\sqrt{2\ln n}\}$, 有

$$\mathbf{P}(A_n) \sim \frac{n^{-c^2}}{c\sqrt{4\pi\ln n}}.$$

由此, 并根据博雷尔–坎泰利引理 (当 $c > 1$ 时, 有 $\sum \mathbf{P}(A_n) < \infty$; 而当 $0 < c < 1$ 时, 有 $\sum \mathbf{P}(A_n) = \infty$), 以及 (3) 和 (4) 的蕴涵关系, 即可推出所证的命题.

2. 设 $\xi_1, \xi_2, \cdots$ 为独立随机变量序列, 均服从参数为 $\lambda > 0$ 的泊松分布. 证明, (与 λ 的值无关)

$$\mathbf{P}\left\{\overline{\lim_n} \frac{\xi_n \ln\ln n}{\ln n} = 1\right\} = 1.$$

提示: 考察事件 $A_n = \{\xi_n > c\varphi_n\}$, 其中 $c > 0$, 而 $\varphi_n = \frac{\ln n}{\ln\ln n}$. 则当 $c > 1$ 时, 有 $\sum \mathbf{P}(A_n) < \infty$; 而当 $0 < c < 1$ 时, 有 $\sum \mathbf{P}(A_n) = \infty$. 再利用博雷尔–坎泰利引理, 以及 (3) 和 (4) 的蕴涵关系.

3. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布的随机变量序列, 具有特征函数

$$\mathbf{E}\, e^{it\xi_1} = e^{-|t|^\alpha}, \quad 0 < \alpha < 2$$

(参阅第三章第 6 节第 4 小节). 证明,

$$\mathbf{P}\left\{\overline{\lim_n} \left|\frac{S_n}{n^{1/\alpha}}\right| = e^{1/\alpha}\right\} = 1.$$

4. 设 $\xi_1, \xi_2, \cdots$ 为伯努利随机变量序列, 有 $\mathbf{P}\{\xi_n = \pm 1\} = 1/2$. 令 $S_n = \xi_1 + \cdots + \xi_n$. 证明如下的哈代–李特尔伍德 (Hardy-Littlewood) 的结论: 概率为 1 地有

$$\lim_{n\to\infty} \frac{|S_n|}{\sqrt{2n\ \ln n}} \leqslant 1.$$

提示: 证明, 对 $a > 0$ 和 $h > 0$, 有

$$\mathbf{P}\{S_n \geqslant a\} \leqslant e^{-ha}\mathbf{E}\, e^{hS_n}$$

及 $\operatorname{ch} h \leqslant \exp\left\{\frac{h^2}{2}\right\}$, 推出不等式

$$\mathbf{P}\{S_n \geqslant a\} \leqslant \exp\left\{-\frac{a^2}{2n}\right\}.$$

再令 $a = 1 + \varepsilon$, $\varepsilon > 0$, 并利用博雷尔–坎泰利引理. (亦可参阅《概率》第二卷 p335 和 p337 中的图书文献资料.)

5. 验证对于不等式 (9) 的如下形式的推广的正确性: 设 $\xi_1, \cdots, \xi_n$ 为相互独立的随机变量, $S_0 = 0$, $S_k = \xi_1 + \cdots + \xi_k$, $k \leqslant n$. 则对任何实数 a, 都有如下的莱维 (Levy) 不等式成立:

$$\mathbf{P}\left\{\max_{0\leqslant k\leqslant n} [S_k + \mu(S_n - S_k)] > a\right\} \leqslant 2\mathbf{P}\{S_n > a\},$$

其中, $\mu(\xi)$ 是随机变量 ξ 的中位数, 即满足如下条件的常数:

$$\max\left(\mathbf{P}\{\xi > \mu(\xi)\},\ \mathbf{P}\{\xi < \mu(\xi)\}\right) \leqslant \frac{1}{2}.$$

(关于中位数的各种不同形式的定义可见第一章第 4 节第 23 题.)

提示:

$$\tau = \inf\{0 \leqslant k \leqslant n:\ S_k + \mu(S_n - S_k) > a\},$$

其中, $\inf\varnothing = n + 1$, 且

$$\mathbf{P}\{S_n > a\} \geqslant \frac{1}{2}\sum_{k=0}^{n} \mathbf{P}\{\tau = k\} = \frac{1}{2}\mathbf{P}\left\{\max_{0\leqslant k\leqslant n} [S_k - \mu(S_n - S_k)] > a\right\}.$$

6. 设 $\xi_1, \cdots, \xi_n$ 为相互独立的随机变量, 有 $\mathbf{E}\,\xi_i = 0$, $1 \leqslant i \leqslant n$. 令 $S_k = \xi_1 + \cdots + \xi_k$. 证明, 对 $a > 0$, 有

$$\mathbf{P}\left\{\max_{1\leqslant k\leqslant n} S_k > a\right\} \leqslant 2\mathbf{P}\{S_n > \varepsilon - \mathbf{E}\,|S_n|\}.$$

7. 设 $\xi_1, \cdots, \xi_n$ 为独立同分布的随机变量, 有 $\mathbf{E}\xi_1 = 0$, $\sigma^2 = \mathbf{E}\xi_1^2 < \infty$ 和 $|\xi_1| \leqslant c$ ($\mathbf{P}$-a.s.), $i \leqslant n$. 令 $S_n = \xi_1 + \cdots + \xi_n$. 证明, 对一切 $0 \leqslant x \leqslant 2c^{-1}$, 都有

$$\mathbf{E}e^{xS_n} \leqslant \exp\left\{\frac{nx^2\sigma^2(1 + xc)}{2}\right\}.$$

在上述条件之下, 再设存在实数序列 (a_n), 当 $n \to \infty$ 时, 有 $a_n/\sqrt{n} \to \infty$ 和 $a_n = o(n)$. 证明, 对一切 $\varepsilon > 0$ 和充分大的 n, 都有

$$\mathbf{P}\{S_n > a_n\} > \exp\left\{-\frac{a_n^2}{2n\sigma^2}(1 + \varepsilon)\right\}.$$

8. 设 $\xi_1, \cdots, \xi_n$ 为独立同分布的随机变量, 有 $\mathbf{E}\xi_i = 0$, $|\xi_i| \leqslant c$ ($\mathbf{P}$-a.s.), $i \leqslant n$. 令 $D_n = \sum\limits_{i=1}^{n} \mathbf{D}\xi_i$. 证明, 对于 $S_n = \xi_1 + \cdots + \xi_n$, 有如下的普罗霍洛夫不等式成立:

$$\mathbf{P}\{S_n > a\} \leqslant \exp\left\{-\frac{a}{2c}\arcsin\frac{ac}{2D_n}\right\}, \quad a \in \mathbb{R}.$$

9. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布的随机变量序列, 对某 $0 < \alpha < 2$, 有 $\mathbf{E}|\xi_1|^\alpha = \infty$. 证明, ($\mathbf{P}$-a.s.地) 有

$$\overline{\lim_n} \frac{|S_n|}{n^{1/\alpha}} = \infty.$$

(由此重对数律不成立.)

10. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布的随机变量序列, 有 $\mathbf{E}\xi_n = 0$, $\mathbf{E}\xi_n^2 = 1$. 令 $S_n = \xi_1 + \cdots + \xi_n$, $n \geqslant 1$. 证明, 随机变量序列

$$\left(\frac{S_n}{\sqrt{2n\ln\ln n}}\right)_{n\geqslant 1}$$

的极限点集概率为 1 地重合于区间 $[-1, 1]$.

11. 设 $\xi_1, \xi_2, \cdots$ 为相互独立的随机变量序列, 均服从正态分布 $\mathcal{N}(m, \sigma^2)$. 令

$$\widetilde{m}_n = \frac{1}{n}\sum_{i=1}^{n}\xi_i.$$

试由上题的结果推出: 随机变量序列

$$\left(\sqrt{n}\frac{\widetilde{m}_n - m}{\sqrt{2n\ln\ln n}}\right)_{n\geqslant 1}$$

的极限点集概率为 1 地重合于区间 $[-\sigma, \sigma]$.

12. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布的随机变量序列, 具有连续的分布函数 $F(x)$, $x \in \mathbb{R}$. 记

$$F_n(x;\omega) = \frac{1}{n}\sum_{k=1}^{n} I(\xi_k(\omega) \leqslant x), \quad x \in \mathbb{R}$$

为经验分布函数, $n \geqslant 1$.

证明, 概率为 1 地有

$$\varlimsup_n \frac{\sqrt{n}\sup_x |F_n(x;\omega) - F(x)|}{\sqrt{2n\ln\ln n}} = \sup_x \sqrt{F(x)(1-F(x))}.$$

13. 设 $\xi_1, \xi_2, \cdots$ 为独立随机变量序列, 具有指数分布 $\mathbf{P}\{\xi_i > x\} = e^{-x}$, $x \geqslant 0$. 试以博雷尔-坎泰利引理中的断言作为基础 (亦可参阅第 1 节第 20 题), 证明, 概率为 1 地有

$$\varlimsup_n \frac{\xi_n}{\ln n} = 1, \quad \varlimsup_n \frac{\xi_n - \ln n}{\ln\ln n} = 1, \quad \varlimsup_n \frac{\xi_n - \ln n - \ln\ln n}{\ln\ln\ln n} = 1.$$

如果随机变量的分布为 $\mathbf{P}\{\xi_i > x\} = e^{-\lambda x}$, $x \geqslant 0$, 其中 $\lambda > 0$, 那么上述结果有什么可以看得见的变化?

14. 设 $\xi_1, \xi_2, \cdots$ 为独立随机变量序列, 具有指数分布 $\mathbf{P}\{\xi_i > x\} = e^{-\lambda x}$, $x \geqslant 0$, 其中 $\lambda > 0$. 证明, 如果 $M_n = \max(\xi_1, \cdots, \xi_n)$, 则有

$$\varlimsup_n \frac{M_n}{\lambda\ln n} = \varlimsup_n \frac{\xi_n}{\lambda\ln n} \quad (\mathbf{P}\text{-a.s.}).$$

15. 设 $\xi_1, \cdots, \xi_n$ 为相互独立的随机变量, 记 $S_0 = 0$, $S_k = \xi_1 + \cdots + \xi_k$, $k \leqslant n$. 证明,

(a) (补充第 5 题) 我们有

$$\mathbf{P}\left\{\max_{1\leqslant k\leqslant n}|S_k + \mu(S_n - S_k)| \geqslant a\right\} \leqslant 2\mathbf{P}\{|S_n| \geqslant a\},$$

其中, $\mu(\xi)$ 是随机变量 ξ 的中位数.

(b) 如果 $\xi_1, \cdots, \xi_n$ 同分布并且对称, 则有

$$1 - e^{-n\mathbf{P}\{|\xi_1|>x\}} \leqslant \mathbf{P}\left\{\max_{1\leqslant k\leqslant n}|\xi_k| > x\right\} \leqslant 2\mathbf{P}\{|S_n| > x\}.$$

16. 设 $\xi_1, \cdots, \xi_n$ 为相互独立的随机变量, 记 $S_k = \xi_1 + \cdots + \xi_k$, $1 \leqslant k \leqslant n$. 证明如下的斯科罗霍德不等式: 对任何 $\varepsilon > 0$, 都有

$$\mathbf{P}\left\{\max_{1\leqslant k\leqslant n}|S_k| \geqslant 2\varepsilon\right\} \leqslant \min_{1\leqslant k\leqslant n}\mathbf{P}\{|S_n - S_k| < \varepsilon\} \cdot \mathbf{P}\{|S_n| \geqslant \varepsilon\}.$$

提示: 考察停时 $\tau = \inf\{1 \leqslant k \leqslant n :\ |S_k| \geqslant 2\varepsilon\}$(令 $\inf\varnothing = n+1$), 并采用第 5 题中的证明思路.

17. 设 $\xi_1, \cdots, \xi_n$ 为随机变量, 记 $S_k = \xi_1 + \cdots + \xi_k$, $1 \leqslant k \leqslant n$. 证明, 对任何 $\varepsilon > 0$, 都有

$$\mathbf{P}\left\{\max_{1\leqslant k\leqslant n}|\xi_k| \geqslant \varepsilon\right\} \leqslant 2\mathbf{P}\left\{\max_{1\leqslant k\leqslant n}|S_k| \geqslant \frac{\varepsilon}{2}\right\}.$$

而如果随机变量 $\xi_1, \cdots, \xi_n$ 相互独立, 并且都服从对称分布, 则对任何 $\varepsilon > 0$, 都有

$$\mathbf{P}\left\{\max_{1\leqslant k\leqslant n}|\xi_k| \geqslant \varepsilon\right\} \leqslant 2\mathbf{P}\left\{|S_n| \geqslant \frac{\varepsilon}{2}\right\}.$$

§5. 强大数定律的收敛速度和大偏差概率

1. 试证明不等式 (8) 与 (20).

提示: 令 $\widetilde{\xi} = -\xi$, 首先证明

$$\widetilde{H}(a) = \sup_{\lambda\in\mathbb{R}}[a\lambda - \psi(\lambda)] = H(-a).$$

再利用不等式 (7).

2. 证明, 在集合 Λ(参阅 (5)) 的内部, 函数 $\psi(\lambda)$ 是下凸的 (如果随机变量 ξ 非退化, 则是严格凸的), 并且无限次可微.

提示: 令 $\lambda_* = \inf_{\lambda\in\Lambda}\lambda$, $\lambda^* = \sup_{\lambda\in\Lambda}\lambda$, 证明 (在条件 (3) 之下, 有)

$$-\infty \leqslant \lambda_* < 0 < \lambda^* \leqslant \infty,$$

并且在区间 (λ_*, λ^*) 上, 函数 $\varphi(\lambda) = \mathbf{E}e^{\lambda\xi}$ 无限次可微. 至于 $\psi(\lambda) = \ln\varphi(\lambda)$ 的凸性, 可由赫尔德不等式得出.

3. 在随机变量 ξ 非退化的假定之下, 证明 $H(a)$ 在整个实直线上可微, 并且是 (下) 凸的.

提示: 确认

$$H(a) = \begin{cases} \lambda_* a - \psi(\lambda_*), & \text{如果 } a \leqslant a_*, \\ a\lambda_0(a) - \psi(\lambda_0(a)), & \text{如果 } a_* < a < a^*, \\ \lambda^* a - \psi(\lambda^*), & \text{如果 } a \geqslant a^*. \end{cases}$$

其中, $\psi(\lambda)=\ln\varphi(\lambda)$, 而

$$a_*=\lim_{\lambda\downarrow\lambda_*}\psi'(\lambda),\quad a^*=\lim_{\lambda\uparrow\lambda^*}\psi'(\lambda).$$

(λ_* 与 λ^* 的定义见上题提示.)

4. 证明如下的关于克拉默变换的反演公式:

$$\psi(\lambda)=\sup_a[\lambda a-H(a)]$$

(对于所有的 λ, 至多去掉集合 $\Lambda=\{\lambda:\ \psi(\lambda)<\infty\}$ 的端点).

5. 设 $\xi_1,\xi_2,\cdots$ 为独立同分布的简单随机变量序列, 有 $\mathbf{E}\xi_1<0$, $\mathbf{P}\{\xi_1>0\}>0$. 记 $S_n=\xi_1+\cdots+\xi_n$, $n\geqslant 1$. 令 $\varphi(\lambda)=\mathbf{E}e^{\lambda\xi_1}$ 和 $\inf\limits_{\lambda}\varphi(\lambda)=p$ $(0<p<1)$.

证明如下结果 (切尔诺夫 (Chernoff) 定理):

$$\lim_n\frac{1}{n}\ln\mathbf{P}\{S_n\geqslant 0\}=\ln p. \tag{$*$}$$

6. 利用 $(*)$ 式证明, 在伯努利场合 ($\mathbf{P}\{\xi_1=1\}=p$, $\mathbf{P}\{\xi_1=0\}=q$), 对 $p<x<1$, 有

$$\lim_n\frac{1}{n}\ln\mathbf{P}\{S_n\geqslant nx\}=-H(x), \tag{$**$}$$

其中 (试比较第一章第 6 节中的记号)

$$H(x)=x\ln\frac{x}{p}+(1-x)\ln\frac{1-x}{1-p}.$$

7. 设 $\xi_1,\xi_2,\cdots$ 为独立同分布的随机变量序列, 有 $\mathbf{E}\xi_1=0$, $\mathbf{D}\xi_1=1$. 令 $S_n=\xi_1+\cdots+\xi_n$, $n\geqslant 1$. 又设 $(x_n)_{n\geqslant 1}$ 为数列, 当 $n\to\infty$ 时, 有 $x_n\to\infty$ 和 $\frac{x_n}{\sqrt{n}}\to 0$.

证明,

$$\mathbf{P}\{S_n\geqslant x_n\sqrt{n}\}=e^{-\frac{x_n^2}{2}(1+y_n)},$$

其中 $y_n\to 0$, $n\to\infty$.

8. 利用 $(**)$ 式证明, 在伯努利场合 ($\mathbf{P}\{\xi_1=1\}=p$, $\mathbf{P}\{\xi_1=0\}=q$), 有:

(a) 对于 $p<x<1$ 和 $x_n=n(x-p)$, 有

$$\mathbf{P}\{S_n\geqslant np+x_n\}=\exp\left\{-nH\left(p+\frac{x_n}{n}\right)(1+o(1))\right\}. \tag{$***$}$$

(b) 对于 $x_n=a_n\sqrt{npq}$, 其中 $a_n\to\infty$, $\frac{a_n}{\sqrt{n}}\to 0$, 有

$$\mathbf{P}\{S_n\geqslant np+x_n\}=\exp\left\{-\frac{x_n^2}{2npq}(1+o(1))\right\}. \tag{$****$}$$

试比较 $(***)$ 式与 $(****)$ 式, 并将它们与第一章第 6 节中的结果进行比较.

9. 设 $\xi_1,\xi_2,\cdots$ 为独立同分布的随机变量序列, 具有柯西分布, 其密度函数为 $f(x)=\frac{1}{\pi(1+x^2)}$, $x\in\mathbb{R}$. 证明,

$$\lim_{n\to\infty}\mathbf{P}\left\{\frac{1}{n}\max_{1\leqslant k\leqslant n}\xi_k<x\right\}=e^{-\frac{1}{\pi x}}.$$

10. 设 $\xi_1,\xi_2,\cdots$ 为独立同分布的随机变量序列, 有 $\mathbf{E}|\xi_1|<\infty$. 证明,

$$\lim_{n\to\infty}\frac{1}{n}\mathbf{E}\left(\max_{1\leqslant k\leqslant n}|\xi_k|\right)=0.$$

(试比较第 3 节第 8 题中的结果.)

11. 设 ξ 为随机变量, 有 $\mathbf{E}\xi=0$ 和 $a\leqslant\xi\leqslant b$. 证明, 对每个 $h>0$, 都有矩母函数

$$\mathbf{E}e^{h\xi}\leqslant e^{\frac{1}{8}h^2(b-a)^2}.$$

提示: e^{hx} 是 x 的下凸函数.

12. 设 $\xi_1,\cdots,\xi_n$ 是独立同分布的伯努利随机变量, 有 $\mathbf{P}\{\xi_1=1\}=p$, $\mathbf{P}\{\xi_1=0\}=q$, $p+q=1$. 记 $S_n=\xi_1+\cdots+\xi_n$. 试证明如下的切尔诺夫不等式: 对于 $x\geqslant 0$, 有

$$\mathbf{P}\{S_n-np\geqslant nx\}\leqslant e^{-2nx^2},$$
$$\mathbf{P}\{|S_n-np|\geqslant nx\}\leqslant 2e^{-2nx^2}.$$

提示: 在这里, 以及在下面的若干习题中, 都应当利用伯恩斯坦不等式:

$$\mathbf{P}\{S_n\geqslant y\}\leqslant e^{-hy}\mathbf{E}e^{hS_n},\quad y\geqslant 0,\ h\geqslant 0.$$

13. 证明, 在上题中的条件下, 事实上还可以有更强的结论成立, 即有如下的关于最大值的不等式:

$$\mathbf{P}\left\{\max_{1\leqslant k\leqslant n}(S_k-kp)\geqslant nx\right\}\leqslant e^{-2nx^2},$$
$$\mathbf{P}\left\{\max_{1\leqslant k\leqslant n}|S_k-kp|\geqslant nx\right\}\leqslant 2e^{-2nx^2}.$$

提示: 利用指数型的柯尔莫戈洛夫不等式

$$\mathbf{P}\left\{\max_{1\leqslant k\leqslant n}(S_k-kp)\geqslant\varepsilon\right\}\leqslant e^{-h\varepsilon}\mathbf{E}e^{h(S_n-np)}$$

(参阅第 2 节第 23 题).

14. 设 $\xi_1,\cdots,\xi_n$ 是相互独立的 (不一定同分布的) 取值于区间 $[0,1]$ 的随机变量. 令 $S_n=\xi_1+\cdots+\xi_n$. 记 $p=\frac{\mathbf{E}S_n}{n}$, $q=1-p$.

证明, 对所有的 $0 \leqslant x < q$, 都有

$$\mathbf{P}\{S_n - \mathbf{E}S_n \geqslant nx\} \leqslant e^{n\psi(x)},$$

其中

$$\psi(x) = \ln\left[\left(\frac{p}{p+x}\right)^{p+x}\left(\frac{q}{q-x}\right)^{q-x}\right].$$

提示: 利用不等式

$$\begin{aligned} e^{hy}\mathbf{P}\{S_n \geqslant y\} &\leqslant \mathbf{E}e^{hS_n} = \mathbf{E}e^{hS_{n-1}}\mathbf{E}e^{h\xi_n} \\ &\leqslant \mathbf{E}e^{hS_{n-1}}(1-p+pe^h) \leqslant \cdots \leqslant (1-p+pe^h)^n, \end{aligned}$$

再选择恰当的 $h > 0$.

15. 在上题中的条件下, 证明如下的霍夫丁不等式, 它是第 12 题中的切尔诺夫不等式的推广: 对于 $x \geqslant 0$, 有

$$\mathbf{P}\{S_n - \mathbf{E}S_n \geqslant nx\} \leqslant e^{-2nx^2},$$
$$\mathbf{P}\{|S_n - \mathbf{E}S_n| \geqslant nx\} \leqslant 2e^{-2nx^2}.$$

提示: 注意到 $\psi(x) \leqslant -2x^2$, 并利用上题中的结果.

16. 设 $\xi_1, \cdots, \xi_n$ 是在区间 $[0, 1]$ 中取值的相互独立的随机变量. 证明, 对任何 $\varepsilon > 0$, 都有如下的不等式成立:

$$\mathbf{P}\Big\{S_n \leqslant (1-\varepsilon)\mathbf{E}S_n\Big\} \leqslant \exp\left\{-\frac{1}{2}\varepsilon^2\mathbf{E}S_n\right\},$$

$$\mathbf{P}\Big\{S_n \geqslant (1+\varepsilon)\mathbf{E}S_n\Big\} \leqslant \exp\Big\{-[(1+\varepsilon)\ln(1+\varepsilon)-\varepsilon]\mathbf{E}S_n\Big\}$$
$$\left(\leqslant \exp\left\{-\frac{\varepsilon^2\mathbf{E}S_n}{2(1+\varepsilon/3)}\right\}\right).$$

提示: 为证前一个不等式, 可利用第 14 题的结果, 注意 $\psi(-xp) \leqslant -px^2/2$, $0 \leqslant x < 1$. 为证第二个不等式, 可利用关于第 14 题的提示, 由它可以得到

$$\mathbf{P}\{S_n - \mathbf{E}S_n \geqslant nx\} \leqslant [e^{(p+x)h}(1-p+pe^h)]^n.$$

17. 设 $\xi_1, \cdots, \xi_n$ 为相互独立的随机变量, 有 $a_i \leqslant \xi_i \leqslant b_i$, 其中 a_i, b_i 为常数, $i = 1, \cdots, n$. 令 $S_n = \xi_1 + \cdots + \xi_n$. 为推广第 15 题中的霍夫丁不等式, 证明, 对 $x \geqslant 0$, 有

$$\mathbf{P}\{S_n - \mathbf{E}S_n \geqslant x\} \leqslant \exp\left\{-2x^2\sum_{k=1}^n (b_k - a_k)^2\right\},$$

$$\mathbf{P}\{|S_n - \mathbf{E}S_n| \geqslant x\} \leqslant 2\exp\left\{-2x^2\sum_{k=1}^n (b_k - a_k)^2\right\}$$

提示: 利用第 11 题中的不等式, 由该式可以推出

$$\mathbf{P}\{S_n - \mathbf{E}S_n \geqslant x\} \leqslant e^{-hx}\mathbf{E}e^{h(S_n - \mathbf{E}S_n)} \leqslant \exp\left\{-hx + \frac{1}{8}h^2\sum_{k=1}^n (b_k - a_k)^2\right\},$$

然后再选择恰当的 h.

18. (大偏差) 设 $(\xi_n)_{n\geqslant 1}$ 是相互独立的标准正态随机变量序列 ($\mathrm{Law}(\xi_n) = \mathcal{N}(0,1)$). 令 $S_n = \xi_1 + \cdots + \xi_n$, $n \geqslant 1$. 证明, 对一切 $A \in \mathscr{B}(\mathbb{R})$, 都有

$$\lim_{n\to\infty}\frac{1}{n}\ln\mathbf{P}\left\{\frac{S_n}{n} \in A\right\} = -\ \text{ess inf}\left\{\frac{x^2}{2}:\ x \in A\right\}.$$

(如果 $f(x)$ 是在 $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ 上给出的实值博雷尔函数, 那么就将 $\operatorname{ess\,inf}\{f(x): x \in A\}$ 理解为 $\sup\Big\{c \in \overline{\mathbb{R}}:\ \lambda\{x \in A:\ f(x) < c\} = 0\Big\}$, 其中 λ 是勒贝格测度; 试比较第二章第 10 节注 3 中关于上确界存在性的定义.)

提示: 试分析表达式

$$\mathbf{P}\left\{\frac{S_n}{n} \in A\right\} = \sqrt{\frac{n}{2\pi}}\int_A e^{-nx^2/2}dx$$

在 $n \to \infty$ 时的极限性状.

19. 设 $\xi = (\xi_1, \cdots, \xi_n)$ 为高斯向量, 有 $\mathbf{E}\xi_i = 0$, $i = 1, \cdots, n$. 证明

$$\lim_{r\to\infty}\frac{1}{\sigma^2}\ln\mathbf{P}\left\{\max_{1\leqslant i\leqslant n}\xi_i \geqslant r\right\} = -\frac{1}{2\sigma^2}.$$

提示: 首先证明, 对一切 $r \geqslant 0$, 都有

$$\mathbf{P}\left\{\max_{1\leqslant i\leqslant n}\xi_i \geqslant \mathbf{E}\max_{1\leqslant i\leqslant n}\xi_i + \sigma r\right\} \leqslant e^{-r^2/2},$$

其中 $\sigma = \max_{1\leqslant i\leqslant n}(\mathbf{E}\xi_i^2)^{1/2}$, 然后再证, 对一切 $1 \leqslant i \leqslant n$, 都有

$$\mathbf{P}\left\{\max_{1\leqslant i\leqslant n}\xi_i \geqslant r\right\} \geqslant 1 - \Phi\left(\frac{r}{\sigma_i}\right) \geqslant \frac{\exp\{-r^2/(2\sigma_i^2)\}}{\sqrt{2\pi}(1 + r/\sigma_i)}.$$

第五章　强 (狭义) 平稳随机序列与遍历理论

§1. 强 (狭义) 平稳随机序列. 保测变换

1. 设 T 是保测变换, $\xi=\xi(\omega)$ 是随机变量, 有数学期望 $\mathbf{E}\xi(\omega)$ 存在. 证明, $\mathbf{E}\xi(\omega)=\mathbf{E}\xi(T\omega)$.

提示: 如果 $\xi=I_A$, $A\in\mathscr{F}$, 则由保测变换的定义知等式 $\mathbf{E}\xi(\omega)=\mathbf{E}\xi(T\omega)$ 成立. 根据线性性质, 知该式对于形如 $\sum\limits_{k=1}^{n}\lambda_k I_{A_k}$, $A_k\in\mathscr{F}$ 的随机变量 ξ 也成立. 接下来就应该利用数学期望的结构和 (关于 $\xi\geqslant 0$ 的) 单调收敛定理. 最后, 对于一般情况, 利用表达式 $\xi=\xi^+-\xi^-$ 即可.

2. 证明, 例 1 和例 2 中的变换 T 是保测变换①.

提示: (关于例 2) 如果 $A=[a,b)\subseteq[0,1)$, 则对这样的 A, 等式 $\mathbf{P}(A)=\mathbf{P}(T^{-1}A)$ 显然成立. 对于一般情况, 令

$$\mathscr{M}=\{A\in\mathscr{B}([0,1]):\ \mathbf{P}(A)=\mathbf{P}(T^{-1}A)\},$$

(运用 "恰当集合" 方法) 证明, $\mathscr{M}=\mathscr{B}([0,1))$.

3. 设 $\Omega=[0,1)$, $\mathscr{F}=\mathscr{B}([0,1))$, 而 $\mathbf{P}$ 是某个具有连续分布函数的概率测度. 证明, 变换 $Tx=\lambda x$, $0<\lambda<1$, 和变换 $Tx=x^2$ 都不是保测变换.

提示: 由于测度 $\mathbf{P}$ 的分布函数连续, 所以可以找到 $a,b\in(0,1)$, 使得

$$\mathbf{P}([0,a))=\frac{1}{3},\quad \mathbf{P}([0,b))=\frac{2}{3}.$$

①在本章中, 一律假定概率空间 $(\Omega,\mathscr{F},\mathbf{P})$ 是完备的.

利用这一性质, 即可证明 $Tx=\lambda x\,(0<\lambda<1)$ 与 $Tx=x^2$ 都不是保测变换.

4. 设 Ω 是所有两端无限的实数序列 $\omega=(\cdots,\omega_{-1},\omega_0,\omega_1,\cdots)$ 的集合, $\mathscr{F}$ 是由可测柱集 $\{\omega:\ (\omega_k,\cdots,\omega_{k+n-1})\}\in B_n$, 其中 $n=1,2,\cdots;\ k=0,\pm1,\pm2,\cdots$, $B_n\in\mathscr{B}(\mathbb{R}^n)$, 所生成的 σ-代数. 设 $\mathbf{P}$ 是可测空间 $(\Omega,\mathscr{F})$ 上的概率测度. 定义两端变换 T 为

$$T(\cdots,\omega_{-1},\omega_0,\omega_1,\cdots)=(\cdots,\omega_0,\omega_1,\omega_2,\cdots).$$

证明, T 是保测变换, 当且仅当, 对一切 $n=1,2,\cdots;\ k=0,\pm1,\pm2,\cdots$ 和 $B_n\in\mathscr{B}(\mathbb{R}^n)$, 都有

$$\mathbf{P}\{\omega:\ (\omega_0,\cdots,\omega_{n-1})\in B_n\}=\mathbf{P}\{\omega:\ (\omega_k,\cdots,\omega_{k+n-1})\in B_n\}.$$

5. 设 $\xi_0,\xi_1,\cdots$ 是取值于某个博雷尔空间 $\mathcal{S}$ 的平稳随机元序列 (参阅第二章第 7 节定义 9). 证明, 可以 (允许是在增广后的原概率空间上) 构造出一个也取值于 $\mathcal{S}$ 的随机元序列 $\xi_{-1},\xi_{-2},\cdots$, 使得两端无限的序列 $\cdots,\xi_{-1},\xi_0,\xi_1,\cdots$ 是平稳的.

6. 设 T 是概率空间 $(\Omega,\mathscr{F},\mathbf{P})$ 上的可测变换, $\mathscr{E}$ 是 Ω 中的子集所组成的生成 $\mathscr{F}$ 的 π-系 (即 $\pi(\mathscr{E})=\mathscr{F}$). 证明, 如果等式 $\mathbf{P}(T^{-1}A)=\mathbf{P}(A)$ 对所有 $A\in\mathscr{E}$ 成立, 则它对所有 $A\in\mathscr{F}$ 都成立.

7. 设 T 是概率空间 $(\Omega,\mathscr{F},\mathbf{P})$ 上的保测变换, $\mathscr{G}$ 是 $\mathscr{F}$ 的子 σ-代数. 证明, 对于每个 $A\in\mathscr{F}$, 都有

$$\mathbf{P}(A\,|\,\mathscr{G})(T\omega)=\mathbf{P}(T^{-1}A\,|\,T^{-1}\mathscr{G})(\omega)\quad(\mathbf{P}\text{-a.s.}).\qquad(*)$$

特别地, 如果 $\Omega=\mathbb{R}^\infty$, 即由数列 $\omega=(\omega_0,\omega_1,\omega_2,\cdots)$ 所构成的空间, 而 $\xi_k(\omega)=\omega_k$. 设 T 是移步变换: $T(\omega_0,\omega_1,\cdots)=(\omega_1,\omega_2,\cdots)$(换言之, 如果 $\xi_k(\omega)=\omega_k$, 则 $\xi_k(T\omega)=\omega_{k+1}$). 那么, 此时 $(*)$ 式具有如下形式:

$$\mathbf{P}(A\,|\,\xi_n)(T\omega)=\mathbf{P}(T^{-1}A\,|\,\xi_{n+1})(\omega)\quad(\mathbf{P}\text{-a.s.}).$$

8. 设 T 是 $(\Omega,\mathscr{F})$ 中的可测变换, $\mathscr{P}$ 是 T 所能保持测度 $\mathbf{P}$ 不变的所有概率测度 $\mathbf{P}$ 的集合. 证明,

(a) 集合 $\mathscr{P}$ 是凸的.

(b) T 关于测度 $\mathbf{P}$ 是遍历的, 当且仅当, $\mathbf{P}$ 是集合 $\mathscr{P}$ 的边缘上的点 (亦即, $\mathbf{P}$ 不能表示为 $\mathbf{P}=\lambda_1\mathbf{P}_1+\lambda_2\mathbf{P}_2$ 的形式, 其中 $\lambda_1>0$, $\lambda_2>0$, $\lambda_1+\lambda_2=1$, $\mathbf{P}_1\neq\mathbf{P}_2$, $\mathbf{P}_1,\mathbf{P}_2\in\mathscr{P}$).

9. 设 T 是概率空间 $(\Omega,\mathscr{F},\mathbf{P})$ 上的保测变换, 而 $\xi=\xi(\omega)$ 是某个随机变量. 证明, $\xi=\xi(\omega)$ 是几乎不变的随机变量 (亦即 $\xi(\omega)=\xi(T\omega)$ ($\mathbf{P}$-a.s.)), 当且仅当, 对每

个有界的 $\mathscr{F}\otimes\mathscr{B}(\mathbb{R})$ 可测的函数 $G(\omega, x)$, 都有

$$\mathbf{E}G(\omega,\xi(\omega))=\mathbf{E}G(T\omega,\xi(\omega)).$$

提示: 首先考察形如 $G_1(\omega)G_2(x)$ 的函数 $G(\omega, x)$.

§2. 遍历性与混合性

1. 证明, 随机变量 ξ 是不变的, 当且仅当, 它是 $\mathscr{J}$ 可测的.

2. 证明, 集合 A 是几乎不变的, 当且仅当, $\mathbf{P}(T^{-1}A\backslash A)=0$.

证明, 如果 X 是几乎不变的随机变量 (即 $X(\omega)=X(T\omega)$ a.s.), 则存在不变的随机变量 $\widetilde{X}=\widetilde{X}(\omega)$(亦即对一切 $\omega\in\Omega$, 都有 $\widetilde{X}(\omega)=\widetilde{X}(T\omega)$), 使得 $\mathbf{P}\{X(\omega)=\widetilde{X}(\omega)\}=1$.

3. 证明, 变换 T 是混合的, 当且仅当, 对任何满足条件 $\mathbf{E}\xi^2<\infty$ 与 $\mathbf{E}\eta^2<\infty$ 的随机变量 ξ 与 η, 都有

$$\mathbf{E}\xi(T^n\omega)\eta(\omega)\to\mathbf{E}\xi(\omega)\mathbf{E}\eta(\omega),\quad n\to\infty. \tag{*}$$

提示: 如果 $\eta=I_A$, $\xi=I_B$, 则性质 (*) 就是混合的性质. 而 L^2 中的每个随机变量 ξ 与 η 都可以 (在 L^2 中的距离之下) 用示性函数的线性组合逼近到精确度为任意的 $\varepsilon>0$ 的程度. 有鉴于此, 不难由混合的性质推出所要证明的数学期望的收敛性.

4. 试举例说明, 存在保测的遍历变换不是混合的.

提示: 考察 $\Omega=\{a,b\}$, $\mathbf{P}(\{a\})=\mathbf{P}(\{b\})=1/2$, 而变换 T 则定义为 $Ta=b$, $Tb=a$.

5. 设 T 是概率空间 $(\Omega,\mathscr{F},\mathbf{P})$ 上的保测变换. 设 $\mathscr{A}$ 是 Ω 的某个子集类, 有 $\sigma(\mathscr{A})=\mathscr{F}$. 假设定义 4 中的关系式

$$\lim_{n\to\infty}\mathbf{P}(A\cap T^{-n}B)=\mathbf{P}(A)\mathbf{P}(B)$$

仅对 $A, B\in\mathscr{A}$ 成立. 证明, 该性质对 $\mathscr{F}=\sigma(\mathscr{A})$ 中的所有 A 与 B 都成立.

再证明, 如果 $\mathscr{A}$ 是使得 $\mathscr{F}=\pi(\mathscr{A})$ 的 π-系, 则上述断言仍然成立.

6. 设 $(\Omega,\mathscr{F})=(\mathbb{R}^\infty,\ \mathscr{B}(\mathbb{R}^\infty))$, 而对于 $\omega=(x_1,x_2,\cdots)$, 变换 T 是移步变换: $T(x_1,x_2,\cdots)=(x_2,x_3,\cdots)$. 证明, 每一个不变集合都是 "尾的" (换言之, 不变集合的 σ-代数 $\mathscr{J}$ 属于 "尾" σ-代数 $\mathscr{X}=\bigcap\mathscr{F}_n^\infty$, 其中 $\mathscr{F}_n^\infty=\sigma(\omega:\ x_n,x_{n+1},\cdots)$). 试给出不是不变的 "尾" 事件的例子.

7. 试分别给出 $(\Omega,\mathscr{F},\mathbf{P})$ 上的保测变换 T 的例子, 使得

(a) 由 $A\in\mathscr{F}$ 根本推不出 $TA\in\mathscr{F}$.

(b) 由 $A\in\mathscr{F}$ 与 $TA\in\mathscr{F}$ 根本推不出 $\mathbf{P}(A)=\mathbf{P}(TA)$.

§3. 遍历性定理

1. 设 $\xi=(\xi_1,\xi_2,\cdots)$ 是平稳的高斯序列, 有 $\mathbf{E}\xi_n=0$ 和协方差函数 $R(n)=\mathbf{E}\xi_{k+n}\xi_k$. 证明, 条件 $R(n)\to 0$ 是使得与 ξ 相应的保测变换为混合的 (因而也是遍历的) 的充分条件.

提示: 如果 $A=\{\omega:\ (\xi_1,\xi_2,\cdots)\in A_0\}$, $B=\{\omega:\ (\xi_1,\xi_2,\cdots)\in B_0\}$, 而 $B_n=\{\omega:\ (\xi_n,\xi_{n+1},\cdots)\in B_0\}$, 则应当证明

$$\mathbf{P}(A\cap B_n)\to\mathbf{P}(A)\mathbf{P}(B),\quad n\to\infty.$$

该性质的证明步骤如下:

(i) (根据给定的 $\varepsilon>0$,) 找出 $m\in\mathbb{N}=\{1,2,\cdots\}$ 和集合 $\widetilde{A}_0\in\mathscr{B}(\mathbb{R}^m)$, $\widetilde{B}_0\in\mathscr{B}(\mathbb{R}^m)$, 使得

$$\mathbf{P}(A\triangle\widetilde{A})<\varepsilon,\quad \mathbf{P}(B\triangle\widetilde{B})<\varepsilon,$$

其中 $\widetilde{A}=\{\omega:\ (\xi_1,\cdots,\xi_m)\in\widetilde{A}_0\}$, $\widetilde{B}=\{\omega:\ (\xi_1,\cdots,\xi_m)\in\widetilde{B}_0\}$.

(ii) 再找出开集 $\overline{A}_0\in\mathscr{B}(\mathbb{R}^m)$ 和 $\overline{B}_0\in\mathscr{B}(\mathbb{R}^m)$, 使得集合

$$\overline{A}=\{\omega:\ (\xi_1,\cdots,\xi_m)\in\overline{A}_0\}\quad 和\quad \overline{B}=\{\omega:\ (\xi_1,\cdots,\xi_m)\in\overline{B}_0\}$$

满足条件

$$\mathbf{P}(\overline{A}\triangle\widetilde{A})<\varepsilon,\quad \mathbf{P}(\overline{B}\triangle\widetilde{B})<\varepsilon,$$

从而就有 $\mathbf{P}(\overline{A}\triangle A)<2\varepsilon$, $\mathbf{P}(\overline{B}\triangle B)<2\varepsilon$.

(iii) 由平稳性知, 对于集合 $\overline{B}_n=\{\omega:\ (\xi_n,\cdots,\xi_{n+m-1})\in\overline{B}_0\}$, 也有 $\mathbf{P}(\overline{B}_n\triangle B_n)<2\varepsilon$.

(iv) 设 P 是随机向量 $(\xi_1,\cdots,\xi_m)$ 的分布, 而 Q_n 是随机向量 $(\xi_1,\cdots,\xi_m, \xi_n,\cdots,\xi_{n+m-1})$ 的分布, 则有

$$R(n)\to 0\quad\Rightarrow\quad Q_n\xrightarrow{\mathrm{w}}P\otimes P,\quad n\to\infty.$$

(v) 根据第三章第 1 节定理 1, 可以由 (iv) 得出

$$\varliminf_n\mathbf{P}(\overline{A}\triangle\overline{B}_n)\geqslant\mathbf{P}(\overline{A})\mathbf{P}(\overline{B}).$$

结合已证的关系式 $\mathbf{P}(\overline{A}\triangle A)<2\varepsilon$ 与 $\mathbf{P}(\overline{B}_n\triangle B_n)<2\varepsilon$, 可得

$$\varliminf_n\mathbf{P}(A\cap B_n)\geqslant(\mathbf{P}(A)-2\varepsilon)(\mathbf{P}(B)-2\varepsilon)-4\varepsilon,$$

由此并由 $\varepsilon>0$ 的任意性, 即得

$$\varliminf_n\mathbf{P}(A\cap B_n)\geqslant\mathbf{P}(A)\mathbf{P}(B).$$

(vi) 可类似证得

$$\overline{\lim_n}\mathbf{P}(A\cap B_n)\leqslant \mathbf{P}(A)\mathbf{P}(B).$$

(不过, 应当将选取开集 $\overline{A}_0$ 和 $\overline{B}_0$ 改为选取闭集.)

2. 证明, 对于任何由独立同分布的随机变量所组成的序列 $\xi=(\xi_1,\xi_2,\cdots)$, 与其相应的保测变换都是混合的.

提示: (首先指出, 在现在的情况下, 序列的遍历性可以直接由 "$0-1$ 律" 推出.) 为证混合性质, 可按下述步骤进行:

(i) 令

$$A=\{\omega:\ (\xi_1,\xi_2,\cdots)\in A_0\},\quad B=\{\omega:\ (\xi_1,\xi_2,\cdots)\in B_0\},$$

其中 $A_0,B_0\in\mathscr{B}(\mathbb{R}^\infty)$. 根据所取的 $\varepsilon>0$, 找出 $m\in\mathbb{N}=\{1,2,\cdots\}$ 和集合 $\widetilde{A}_0\in\mathscr{B}(\mathbb{R}^m)$, 使得对于集合 $\widetilde{A}=\{\omega:\ (\xi_1,\cdots,\xi_m)\in\widetilde{A}_0\}$, 有 $\mathbf{P}(A\triangle\widetilde{A})<\varepsilon$.

(ii) 再令

$$B_n=\{\omega:\ (\xi_n,\xi_{n+1},\cdots)\in B_0\}.$$

那么, 当 $n>m$ 时, 就有

$$\mathbf{P}(\widetilde{A}\cap B_n)=\mathbf{P}(\widetilde{A})\mathbf{P}(B_n)=\mathbf{P}(\widetilde{A})\mathbf{P}(B).$$

(iii) 试推出 $|\mathbf{P}(A\cap B_n)-\mathbf{P}(A)\mathbf{P}(B_n)|\leqslant 2\varepsilon$, 以及当 $n\to\infty$ 时, 有 $\mathbf{P}(A\cap B_n)\to\mathbf{P}(A)\mathbf{P}(B_n)$.

3. 证明, 平稳序列 ξ 是遍历的, 当且仅当, 对任何 $B\in\mathscr{B}(\mathbb{R}^k)$, $k=1,2,\cdots$, 都有

$$\frac{1}{n}\sum_{i=1}^{n}I_B(\xi_i,\cdots,\xi_{i+k-1})\to\mathbf{P}\{(\xi_1,\cdots,\xi_k)\in B\}\quad(\mathbf{P}\text{-a.s.}).$$

提示: 为证必要性, 我们以 Q 表示序列 $\xi=(\xi_1,\xi_2,\cdots)$ 的分布 (Q 是 $\mathbb{R}^\infty$ 上的测度). 令 T 为移步变换:

$$T:\quad x=(x_1,x_2,\cdots)\in\mathbb{R}^\infty\to x'=(x_2,x_3,\cdots)\in\mathbb{R}^\infty.$$

还应该对 $k=1,2,\cdots$, 引入函数 $f:\ (x_1,x_2,\cdots)\in\mathbb{R}^\infty\to I((x_1,\cdots,x_k)\in B)\in\mathbb{R}$, 并将遍历定理 (伯克霍夫–辛钦定理) 用于该函数; $B\in\mathscr{B}(\mathbb{R}^k)$.

为证充分性, 应当证明所引入的变换 T 是遍历的, 换言之, $\mathscr{J}$ 中的每个集合 (即不变集合) 的测度都是 0 或 1.

性质

$$\frac{1}{n}\sum_{i=1}^{n}I((\xi_1,\cdots,\xi_k)\in B)\to\mathbf{P}\{(\xi_1,\cdots,\xi_k)\in B\}\quad(\mathbf{P}\text{-a.s.})$$

意味着对每个形如 $\{(x_1,\cdots,x_k)\in B\}$, $B\in\mathscr{B}(\mathbb{R}^k)$ 的集合 $A\in\mathscr{B}(\mathbb{R}^\infty)$, 都有

$$\frac{1}{n}\sum_{i=1}^{n}I_A(T^ix)\to Q(A)\quad(Q\text{-a.s.}).$$

该式与遍历定理 (伯克霍夫–辛钦定理) 中的断言一起蕴涵 $\mathbf{E}_Q\,(I_A\,|\,\mathscr{J})=\mathbf{E}_QI_A$ (Q-a.s.), 这就表明, 每个形如 $\{(x_1,\cdots,x_k)\in B\}$ 的集合 A 都与 $\mathscr{J}$ 独立, 其中 $B\in\mathscr{B}(\mathbb{R}^k)$. 再运用 "恰当集合" 方法, 证明

$$\mathscr{M}=\{A\in\mathscr{B}(\mathbb{R}^\infty):\ A\text{ 与 }\mathscr{J}\text{ 独立}\}$$

重合于 $\mathscr{B}(\mathbb{R}^\infty)$. 最后, 再由此推出 $\mathscr{J}$ 与 $\mathscr{J}$ 独立, 从而任何不变集合的 Q 测度都是 0 或 1. 这就确立了变换 T 的遍历性, 也就意味着序列 ξ 具有遍历性.

4. 假设在可测空间 $(\Omega,\mathscr{F})$ 上给定了两个概率测度 $\mathbf{P}$ 和 $\overline{\mathbf{P}}$, 关于它们的保测变换 T 是遍历的. 证明, 或者有 $\mathbf{P}=\overline{\mathbf{P}}$, 或者有 $\mathbf{P}\perp\overline{\mathbf{P}}$.

5. 设 T 是概率空间 $(\Omega,\mathscr{F},\mathbf{P})$ 上的保测变换, $\mathscr{A}$ 是 Ω 的某个子集类, 有 $\sigma(\mathscr{A})=\mathscr{F}$. 令

$$I_A^{(n)}=\frac{1}{n}\sum_{k=0}^{n-1}I_A(T^k\omega).$$

证明, 变换 T 是遍历的, 当且仅当, 下列条件之一成立:

(i) 对任何 $A\in\mathscr{A}$, 都有

$$I_A^{(n)}\xrightarrow{\mathbf{P}}\mathbf{P}(A).$$

(ii) 对任何 $A,B\in\mathscr{A}$, 都有

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\mathbf{P}(A\cap T^{-k}B)=\mathbf{P}(A)\mathbf{P}(B).$$

(iii) 对任何 $A\in\mathscr{F}$, 都有

$$I_A^{(n)}\xrightarrow{\mathbf{P}}\mathbf{P}(A).$$

6. 设 T 是概率空间 $(\Omega,\mathscr{F},\mathbf{P})$ 上的保测变换. 证明, 变换 T (关于 $\mathbf{P}$) 是遍历的, 当且仅当, 在 $(\Omega,\mathscr{F})$ 中不存在测度 $\overline{\mathbf{P}}\neq\mathbf{P}$, 使得 $\overline{\mathbf{P}}\ll\mathbf{P}$, 并且关于测度 $\overline{\mathbf{P}}$, T 是保测变换.

7. (伯努利移动) 设 S 是某个有限集合 (记 $S=\{1,2,\cdots,N\}$), 而 $\Omega=S^\infty$ 是由序列 $\omega=(\omega_0,\omega_1,\cdots)$ 所构成的空间, 其中 $\omega_i\in S$. 令 $\xi_k(\omega)=\omega_k$, 并且定义移步变换 $T(\omega_0,\omega_1,\cdots)=(\omega_1,\omega_2,\cdots)$, 如果用 $\xi_0,\xi_1,\cdots$ 来表示, 即如果 $\xi_k(\omega)=\omega_k$, 则有 $\xi_k(T\omega)=\omega_{k+1}$. 假设在集合 $\{1,2,\cdots,N\}$ 中的元素 i 处赋给一个非负实数 p_i, 使得 $\sum\limits_{i=1}^N p_i=1$ (亦即数组 $p_1,p_2,\cdots,p_N$ 形成一个概率分布). 利用这个定义, 可以在空间 $(S^\infty,\ \mathscr{B}(S^\infty))$ 中给出一个测度 $\mathbf{P}$ (参阅第二章第 3 节):

$$\mathbf{P}\{\omega:\ (\omega_1,\cdots,\omega_k)=(u_1,\cdots,u_k)\}=p_{u_1}\cdots p_{u_k}.$$

换言之, 以保证随机变量 $\xi_0(\omega),\xi_1(\omega),\cdots$ 的相互独立性为原则来引入概率测度 $\mathbf{P}$. 关于这个测度 $\mathbf{P}$, 移步变换 T 就称为伯努利移动, 或称为伯努利变换.

证明, 伯努利变换具有混合性质.

8. 设 T 是概率空间 $(\Omega,\mathscr{F},\mathbf{P})$ 上的保测变换. 记 $T^{-n}\mathscr{F}=\{T^{-n}A:\ A\in\mathscr{F}\}$. 我们称 σ-代数

$$\mathscr{F}_{-\infty}=\bigcap_{n=1}^{\infty}T^{-n}\mathscr{F}$$

是平凡的 ($\mathbf{P}$-平凡的), 如果 $\mathscr{F}_{-\infty}$ 中的每个集合的 $\mathbf{P}$ 测度都是 0 或 1. 具有平凡的 σ-代数 $\mathscr{F}_{-\infty}$ 的变换 T 称为柯尔莫戈洛夫变换. 证明, 柯尔莫戈洛夫变换具有遍历性, 并且具有混合性质.

9. 设 $1\leqslant p<\infty$, 而 T 是概率空间 $(\Omega,\mathscr{F},\mathbf{P})$ 上的保测变换. 设随机变量 $\xi(\omega)\in L^p(\Omega,\mathscr{F},\mathbf{P})$.

证明如下的 $L^p(\Omega,\mathscr{F},\mathbf{P})$ 中的冯·诺依曼遍历定理: 存在随机变量 $\eta(\omega)$, 使得

$$\mathbf{E}\left|\frac{1}{n}\sum_{k=0}^{n-1}\xi(T^k\omega)-\eta(\omega)\right|^p\to 0,\quad n\to\infty.$$

10. 关于正则性的博雷尔定理断言 (第四章第 3 节例 2), 区间 $[0,1)$ 中的数 ω 的二进制展开式中的 0 和 1 的比例 (关于勒贝格测度) 几乎处处收敛到 $1/2$. 试通过考察按如下方式所定义的变换 $T:\ [0,1)\to[0,1)$:

$$T(\omega)=2\omega\pmod 1,$$

并利用遍历定理 1, 证明这一断言.

11. 如同第 10 题, 设 $\omega\in[0,1)$. 我们来考察这样的变换 $T:\ [0,1)\to[0,1)$:

$$T(\omega)=\begin{cases}0, & \text{如果 } \omega=0,\\ \left\{\dfrac{1}{\omega}\right\}, & \text{如果 } \omega\neq 0,\end{cases}$$

其中, $\{x\}$ 表示实数 x 的小数部分.

证明, 该变换 T 保持区间 $[0,1)$ 上的按照如下方式定义的高斯测度 $P=P(\cdot)$ 不变:

$$P(A)=\frac{1}{\ln 2}\int_A\frac{dx}{1+x},\quad A\in\mathscr{B}([0,1)).$$

12. 试举例说明, 关于"增长性"的泊松定理 (第 1 节第 3 小节) 在无限维可测空间中可能不成立.

第六章　弱(广义)平稳随机序列. L^2 理论

§1. 协方差函数的谱表示

1. 试由 (11) 推出性质 (12).

提示: 所要求证明的性质可以通过适当选择的数值 t_k 和 a_k, $1 \leqslant k \leqslant n$ 来得到. 例如, 如果取 $m=2$, $t_1=0$, $t_2=n$, 则可得到不等式

$$(|a_1|^2+|a_2|^2)R(0)+a_1\bar{a}_2R(-n)+\bar{a}_1a_2R(n) \geqslant 0.$$

如果在此取 $a_1=a_2=1$ 与 $a_1=1$, $a_2=i$, 并考虑到 $R(0)\in\mathbb{R}=(-\infty,\infty)$, 则可得到 $R(n)+R(-n)\in\mathbb{R}$ 和 $i(R(n)+R(-n))\in\mathbb{R}$, 从而意味着 $R(-n)=\overline{R}(n)$.

2. 证明, 如果由 (27) 式所定义的多项式 $Q(z)$ 的所有 0 点都在单位圆之外, 则自回归方程 (24) 有, 并且只有一个, 平稳解可以表示为单向滑动平均的形式.

3. 证明, (22) 与 (24) 中的谱函数序列分别具有如 (23) 和 (29) 式所给出的密度.

提示: 关于对 (23) 式的证明, 需要按照如下方式进行: 首先断言 $R(n)=\sum\limits_{k=0}^{p}a_{n+k}\bar{a}_k$, 然后验证 $R(n)=\int_{-\pi}^{\pi}e^{i\lambda n}f(\lambda)d\lambda$, 其中 $f(\lambda)$ 由 (23) 式所给出. (采用形式 $\int_{-\pi}^{\pi}e^{i\lambda n}d\lambda=2\pi\delta_{n0}$ 是有益的, 其中 δ_{n0} 是克罗内克符号.)

4. 证明, 如果 $\sum\limits_{n=-\infty}^{\infty}|R(n)|<\infty$, 则谱函数 $F(\lambda)$ 具有如下形式的谱密度 $f(\lambda)$:

$$f(\lambda)=\frac{1}{2\pi}\sum_{n=-\infty}^{\infty}e^{-i\lambda n}R(n),$$

并且该级数在复空间 $L^2=L^2([-\pi,\pi),\ \mathscr{B}([-\pi,\pi),\ \mu)$ 中收敛, 其中 μ 为勒贝格测度.

提示: 利用这样的事实: $\left\{\frac{1}{\sqrt{2\pi}}e^{i\lambda n},\ n=0,\pm1,\pm2,\cdots\right\}$ 是 $L^2=L^2([-\pi,\pi),\ \mathscr{B}([-\pi,\pi),\ \mu)$ 中的正交系, 其中 μ 为勒贝格测度.

5. 设 $(\xi_n)_{n\geqslant0}$ 是具有零均值的高斯–马尔可夫平稳序列. 证明, 协方差函数 $R(n)$ 可以表示为

$$R(n)=\sigma^2\lambda^n,$$

其中 λ 是某个满足 $0<\lambda\leqslant1$ 的实数.

6. 设 $N=(N_t)_{t\geqslant0}$ 是参数为 $\lambda>0$ 的泊松过程 (参阅第七章第 10 节). 定义 (具有连续参数的) 过程 $\xi_t=\xi\cdot(-1)^{N_t}$, 其中 ξ 是与 N 独立的随机变量, 有 $\mathbf{P}\{\xi=1\}=\mathbf{P}\{\xi=-1\}=\frac{1}{2}$. 证明, $\mathbf{E}\xi_t=0$, 且其协方差函数为 $\mathbf{E}\xi_s\xi_t=e^{-2\lambda|t-s|}$, $s,t\geqslant0$.

7. 设 $(\xi_n)_{n\geqslant0}$ 为这样的序列: $\xi_n=\sum\limits_{k=1}^{N}a_k\cos(b_kn-\eta_k)$, 其中, 对一切 $k=1,\cdots,N$, 都有 $a_k>0$, $b_k>0$, 而 $\eta_1,\cdots,\eta_N$ 是独立的随机变量, 且在区间 $(0,2\pi)$ 上均匀分布. 证明, 序列 $(\xi_n)_{n\geqslant0}$ 是广义平稳序列.

8. 设 $\xi_n=\cos n\varphi$, 其中 $n\geqslant1$, 而 φ 服从区间 $[-\pi,\pi]$ 上均匀分布的随机变量. 证明, 序列 $(\xi_n)_{n\geqslant1}$ 是广义平稳序列, 但不是狭义的平稳序列.

9. 我们来考察单向 p 阶滑动平均模型 (MA(p)):

$$\xi_n=a_0\varepsilon_n+a_1\varepsilon_{n-1}+\cdots+a_p\varepsilon_{n-p},$$

其中 $n=0,\pm1,\pm2,\cdots$, 而 $\varepsilon=(\varepsilon_n)$ 是白噪声 (第 1 节例 3). 试求方差 $\mathbf{D}\xi_n$ 和协方差 $\mathbf{cov}\ (\xi_n,\xi_{n+k})$.

10. 我们来考察一阶自回归模型 (AR(1)):

$$\xi_n=\alpha_0+\alpha_1\xi_{n-1}+\sigma\varepsilon_n,\quad n\geqslant1,$$

其中, $\varepsilon=(\varepsilon_n)$ 是白噪声 (试比较第 1 节 (25) 式).

设 $|\alpha_1|<1$. 证明, 如果 $\mathbf{E}|\xi_0|<\infty$, 则有

$$\mathbf{E}\xi_n=\alpha_1^n\mathbf{E}\xi_0+\frac{\alpha_0(1-\alpha_1^n)}{1-\alpha_1}\to\frac{\alpha_0}{1-\alpha_1},\quad n\to\infty;$$

如果还有 $\mathbf{D}|\xi_0|<\infty$, 则有

$$\mathbf{D}\xi_n=\alpha_1^{2n}\mathbf{D}\xi_0+\frac{\sigma^2(1-\alpha_1^{2n})}{1-\alpha_1^2}\to\frac{\sigma^2}{1-\alpha_1^2},\quad n\to\infty;$$

与

$$\mathbf{cov}\ (\xi_n,\xi_{n+k})\to\frac{\sigma^2\alpha_1^k}{1-\alpha_1^2},\quad n\to\infty.$$

11. 设上题中的 ξ_0 服从正态分布 $\mathcal{N}(\frac{\alpha_0}{1-\alpha_1}, \frac{\sigma^2}{1-\alpha_1^2})$. 证明, 序列 $(\xi_n)_{n\geqslant 0}$ 是平稳的 (既是广义的, 又是狭义的) 高斯序列, 并且有

$$\mathbf{E}\xi_n = \frac{\alpha_0}{1-\alpha_1}, \quad \mathbf{D}\xi_n = \frac{\sigma^2}{1-\alpha_1^2}$$

与

$$\mathbf{cov}\,(\xi_n, \xi_{n+k}) = \frac{\sigma^2\alpha_1^k}{1-\alpha_1^2}.①$$

§2. 正交随机测度与随机积分

1. 试证明 (5) 式与 (6) 式的等价性条件.

提示: 为证 (5) $\Rightarrow$ (6), 应当取 $\Delta_n \downarrow \varnothing$, $\Delta_n \in \mathscr{E}_0$, $D_n = E\backslash\Delta_n$, $D_0 = \varnothing$; 于是 $E = \sum_{k=1}^{\infty}(D_k\backslash D_{k-1})$. 由 (5) 推知 $Z(\Delta_n) = Z(E) - Z(D_n) \stackrel{H^2}{\to} 0$ (记号 H^2 见《概率》第二卷 p58).

2. 设 $f \in L^2$, 试利用第二章中的结果 (第 4 节定理 1, 第 6 节定理 3 的推论和第 3 节第 8 题), 证明, 存在形如 (10) 式所示的函数序列 $(f_n)_{n\geqslant 1}$, 使得 $\|f_n - f\| \to 0$, $n \to \infty$.

提示: 证明思路如下: 取 $\varepsilon > 0$, 找简单函数 $g(\lambda) = \sum_{k=1}^{p} f_k I_{B_k}(\lambda)$, 其中 $B_k \in \mathscr{E}$, $f_k \in \mathbb{C}$, 使得 $\|f - g\|_{L^2} < \varepsilon/2$. 然后再找集合 $\Delta_k \in \mathscr{E}_0$, 使得测度 $m(\Delta_k \triangle B_k)$ 足够小, $k = 1, \cdots, p$. 于是函数 $h(\lambda) = \sum_{k=1}^{p} f_k I_{\Delta_k}(\lambda)$ 具有如 (10) 式所示的形式, 并且有 $\|f - h\|_{L^2} < \varepsilon$.

3. 试证明具有函数结构 $m(\Delta)$ 的正交随机测度 $Z(\Delta)$ 的如下性质:

$$\mathbf{E}|Z(\Delta_1) - Z(\Delta_2)|^2 = m(\Delta_1 \triangle \Delta_2),$$
$$Z(\Delta_1\backslash\Delta_2) = Z(\Delta_1) - Z(\Delta_1 \cap \Delta_2) \quad (\mathbf{P}-\text{ a.s.}),$$
$$Z(\Delta_1 \triangle \Delta_2) = Z(\Delta_1) + Z(\Delta_2) - 2Z(\Delta_1 \cap \Delta_2) \quad (\mathbf{P}-\text{ a.s.}).$$

4. 设 $\xi = (\xi_n)$ 是平稳序列, 具有 $\mathbf{E}\xi_n = 0$, 协方差函数 $R(n)$ 和谱测度 $F(d\lambda)$. 令 $S_n = \xi_1 + \cdots + \xi_n$. 证明, 方差 $\mathbf{D}S_n$ 等于

$$\mathbf{D}S_n = \sum_{|k|<n}(n - |k|)R(k) \quad 和 \quad \mathbf{D}S_n = \int_{-\pi}^{\pi} \frac{\sin\frac{n\lambda}{2}}{\sin\frac{\lambda}{2}} F(d\lambda).$$

(其中, 核 $\left(\frac{\sin(n\lambda/2)}{\sin(\lambda/2)}\right)$ 称为费耶尔 (Fejér) 核.)

① 原书为: $\mathbf{cov}\,(\xi_n, \xi_{n+k}) = \frac{\sigma^2\alpha_1^k}{1-\alpha_1^2}$ —— 译者注.

5. 设 $f(\lambda)$ 为谱密度 ($F(d\lambda) = f(\lambda)d\lambda$), 并且 $f(\lambda)$ 在 $\lambda = 0$ 处连续. 试利用上题中的关于 $\mathbf{D}S_n$ 的积分表达式, 证明,

$$\mathbf{D}S_n = 2\pi f(0)\cdot n + o(n), \quad n \to \infty.$$

§3. 弱 (广义) 平稳序列的谱表示

1. 证明, $\overline{L_0^2}(F) = L^2(F)$(有关符号参阅定理 1 的证明).

提示: 根据第 2 节第 2 题的结论, 所有函数 $f(\lambda) \in L^2(F)$ 都可以依 $L^2(F)$ 中的模被形如 $g(\lambda) = \sum_{k=1}^{p} f_k I_{B_k}(\lambda)$ 的函数逼近到任意近的地步, 其中 $B_k \in \mathscr{A}$, 而 $\mathscr{A}$ 是半开闭区间 $[a, b)$ 的有限并的代数, 此处 $-\pi \leqslant a < b < \pi$. 所以只需证明, 形如 $I_{[a,b)}(\lambda)$ 的函数可以用形如 $e_n(\lambda) = e^{i\lambda n}$ 的函数的线性组合来逼近, 其中 $n = 0, \pm1, \pm2, \cdots$. 这些函数可以用连续函数逼近, 而连续函数自身又可以用函数 $e_n(\lambda) = e^{i\lambda n}$, $n = 0, \pm1, \pm2, \cdots$, 的线性组合来逼近 (魏尔斯特拉斯–斯通定理).

2. 设 $\xi = (\xi_n)$ 为平稳序列, 具有如下性质: 对某个 N 和一切 n, 都有 $\xi_{N+n}(\omega) = \xi_n(\omega)$, $\omega \in \Omega$. (此处及下面, 符号 $\xi = (\xi_n)$ 中的 n 均理解为两个无限的整数 $n = 0, \pm1, \pm2, \cdots$.) 证明, 这样的序列的谱表示可以化为第 1 节中 (13) 式.

提示: 由条件 $R(N) = R(0)$ 可以推知, 谱测度 F 在区间 $[-\pi, \pi)$ 中是分段常数的, 其跳跃点为

$$\lambda_k = \frac{2\pi k}{N} + 2\pi p_k, \quad k = 1, \cdots, N,$$

其中 p_k 为整数, 使得 $\lambda_k \in [-\pi, \pi)$, 从而谱表示就呈如下形式:

$$\xi_n = \int_{[-\pi,\pi)} e^{i\lambda n} Z(d\lambda) = \sum_{k=1}^{N} e^{i\lambda_k n} Z(\{\lambda_k\}).$$

3. 设 $\xi = (\xi_n)$ 为平稳序列, 有 $\mathbf{E}\xi_n = 0$, 且对某个 $C > 0$, $\alpha > 0$, 有

$$\frac{1}{N^2}\sum_{k=0}^{N-1}\sum_{l=0}^{N-1} R(k-l) = \frac{1}{N}\sum_{|k|\leqslant N-1} R(k)\left[1 - \frac{|k|}{N}\right] \leqslant CN^{-\alpha}.$$

试利用博雷尔–坎泰利引理, 证明

$$\frac{1}{N}\sum_{k=0}^{N}\xi_k \to 0 \quad (\mathbf{P}\text{-a.s.}), \quad N \to \infty.$$

4. 设序列 $\xi = (\xi_m)$ 的谱密度 $f_\xi(\lambda)$ 为有理的

$$f_\xi(\lambda) = \frac{1}{2\pi}\frac{|P_{n-1}(e^{-i\lambda})|}{|Q_n(e^{-i\lambda})|},$$

其中 $P_{n-1}(z)=a_0+a_1z+\cdots+a_{n-1}z^{n-1}$, 而 $Q_n(z)=1+b_1z+\cdots+b_nz^n$, 并且多项式 Q_n 的根都不在单位圆中.

证明, 存在白噪声 $\varepsilon=(\varepsilon_m)$, 使得序列 (ξ_m) 是 n 维序列 $(\xi_m^1,\cdots,\xi_m^n)$, $\xi_m^1=\xi_m$, 的分量, 满足方程组

$$\xi_{m+1}^i=\xi_m^{i+1}+\beta_i\varepsilon_{m+1},\quad i=1,\cdots,n-1,$$

$$\xi_{m+1}^n=-\sum_{j=0}^{n-1}b_{n-j}\xi_m^{j+1}+\beta_n\varepsilon_{m+1},$$

其中 $\beta_1=a_0$, $\beta_i=a_{i-1}-\sum\limits_{k=1}^{i-1}\beta_kb_{i-k}$.

5. 称狭义平稳序列 $\xi=(\xi_n)$ 满足强混合条件, 如果

$$\alpha_n(\xi)=\sup_{A\in\mathscr{F}_{-\infty}^0(\xi),\ B\in\mathscr{F}_n^\infty(\xi)}|\mathbf{P}(AB)-\mathbf{P}(A)\mathbf{P}(B)|\to 0,\quad n\to\infty,$$

其中 $\mathscr{F}_{-\infty}^0(\xi)=\sigma(\cdots,\xi_{-1},\xi_0)$ 与 $\mathscr{F}_n^\infty=\sigma(\xi_n,\xi_{n+1},\cdots)$ 分别为随机变量族 $(\xi_k)_{k\leqslant 0}$ 与 $(\xi_k)_{k\geqslant n}$ 所生成的 σ-代数. (试比较第二章第 8 节第 7 题.)

证明, 如果 X 与 Y 分别为 σ-代数 $\mathscr{F}_{-\infty}^0(\xi)$ 与 $\mathscr{F}_n^\infty$ 可测的有界随机变量 ($|X|\leqslant C_1$, $|Y|\leqslant C_2$), 则有

$$|\mathbf{E}XY-\mathbf{E}X\mathbf{E}Y|\leqslant 4C_1C_2\,\alpha_n(\xi).$$

6. 设 $\xi=(\xi_m)_{-\infty<m<\infty}$ 是平稳高斯序列, 而

$$\rho_n^*(\xi)=\sup_{X,Y}\mathbf{E}XY,$$

其中的上确界系对所有满足条件 $\mathbf{E}|X|^2=\mathbf{E}|Y|^2=1$ 的分别属于由随机变量族 $(\xi_m)_{m\leqslant 0}$ 与 $(\xi_m)_{m\geqslant n}$ 所生成的闭线性流形 $L_{-\infty}^n(\xi)$ 与 $L_n^\infty(\xi)$ 的所有随机变量 X 与 Y 所取.

证明如下的柯尔莫戈洛夫–罗扎诺夫(Розанов) 不等式

$$\alpha_n(\xi)\leqslant\rho_n^*(\xi)\leqslant 2\pi\alpha_n(\xi).$$

(试比较第二章第 8 节第 7 题中的不等式.)

7. 以上题中的不等式为基础, 证明, 如果平稳的高斯序列 $\xi=(\xi_n)$ 具有连续的谱密度 $f(\lambda)$, 并且该谱密度函数具有一致的正的下界 (即有 $f(\lambda)>C>0$, $\lambda\in[-\pi,\ \pi]$), 则该高斯序列满足强混合条件.

8. 通过恰当选择随机变量 λ 与 θ, 并考察序列 $\xi=(\xi_n)$, 其中 $\xi_n=A\cos(\lambda n+\theta)$, $A\neq 0$, 以确认, 尽管广义平稳序列的协方差函数可能不是周期函数, 但是它却可能具有周期性的样本轨道.

§4. 协方差函数和谱密度的统计估计

1. 设在 (15) 式中, 有 $\xi_n\sim\mathcal{N}(0,1)$. 证明, 对任何 n, 当 $N\to\infty$ 时, 都有

$$(N-|n|)\mathbf{D}\widehat{R}_N(n;\xi)\to 2\pi\int_{-\pi}^{\pi}(1+e^{2in\lambda})f^2(\lambda)d\lambda.$$

提示: 利用 (15) 式中关于高斯变量 ξ_n 的假设条件, 可以证得 ($n\geqslant 0$)

$$(N-|n|)\mathbf{D}\widehat{R}_N(n;\xi)=2\pi\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}(1+e^{in(\lambda+\nu)})\Phi_{N-n}(\lambda-\nu)f(\nu)f(\lambda)d\nu d\lambda,$$

其中 $\Phi_{N-n}(\lambda)$ 是费耶尔核. 由此即可推出所需要的结论.

2. 证明 (16) 式及其如下形式的推广:

$$\lim_{N\to\infty}\mathbf{cov}\,(\widehat{f}_N(\lambda;\xi),\ \widehat{f}_N(\nu;\xi))=\begin{cases}2f^2(0), & \lambda=\nu=0,\pm\pi,\\ f^2(\lambda), & \lambda=\nu\neq 0,\pm\pi,\\ 0, & \lambda\neq\nu.\end{cases}$$

3. 考察一阶自回归模型 (AR(1)):

$$\xi_n=\theta\xi_{n-1}+\sigma\varepsilon_n,\quad n\geqslant 1,$$

(试比较第 1 节 (25) 式和第 10 题模型), 其中, σ 是已知参数 ($\sigma>0$), 而 $\theta\in\mathbb{R}$ 是未知参数, 需要利用观察值 $\xi_1,\xi_2,\cdots$ 对其进行估计.

设 $\widehat{\theta}_n=\arg\max p_\theta(x_1,\cdots,x_n)$ 是参数 θ 的极大似然估计, 其中

$$p_\theta(x_1,\cdots,x_n)=\frac{1}{(\sqrt{2\pi}\sigma)^n}\exp\left\{-\frac{1}{2\sigma^2}\sum_{k=1}^n(x_k-\theta x_{k-1})^2\right\}$$

是随机变量 $\xi_1,\cdots,\xi_n$ 的概率分布密度 (假设 $\varepsilon=(\varepsilon_n)$ 是高斯白噪声).

证明,

$$\widehat{\theta}_n=\frac{\sum\limits_{k=1}^n X_{k-1}X_k}{\sum\limits_{k=1}^n X_{k-1}^2}.$$

4. 设

$$I_n(\theta)=\mathbf{E}_\theta\left\{-\frac{\partial^2\ln p_\theta(\xi_1,\cdots,\xi_n)}{\partial\theta^2}\right\}$$

是第 3 题中模型 (AR(1)) 中的费希尔信息($\mathbf{E}_\theta$ 是根据序列 $\xi_1,\xi_2,\cdots$ 的分布 $\mathbf{P}_\theta$ 计算均值).

证明

(a) $I_n(\theta)=\mathbf{E}_\theta\sum\limits_{k=1}^n\xi_{k-1}^2$.

(b) 当 $n\to\infty$ 时, 有

$$I_n(\theta)\sim\begin{cases}\dfrac{n}{1-\theta^2}, & |\theta|<1,\\ \dfrac{n^2}{2}, & |\theta|=1,\\ \dfrac{\theta^{2n}}{(\theta^2-1)^2}, & |\theta|>1.\end{cases}$$

5. 对于第 3, 4 题中的模型 (AR(1)), 证明, 极大似然估计 $\widehat{\theta}_n$ 具有如下的渐近性质:

$$\lim_{n\to\infty}\mathbf{P}_\theta\{\sqrt{I_n(\theta)}(\widehat{\theta}_n-\theta)\leqslant x\}=\begin{cases}\Phi(x), & |\theta|<1,\\ H_\theta^{(1)}(x), & |\theta|=1,\\ \mathrm{Ch}(x), & |\theta|>1.\end{cases}$$

其中, $\Phi(x)$ 是标准正态分布函数, $H_\theta^{(1)}(x)$ 是如下的随机变量的分布:

$$\theta\cdot\frac{B_1^2-1}{2\sqrt{2}\int_0^1 B_s^2ds}$$

$(B=(B_s)_{0\leqslant s\leqslant 1}$ 是布朗运动, 见第二章第 13 节), 而 $\mathrm{Ch}(x)$ 是以 $\frac{1}{\pi(1+x^2)}$ 为密度的柯西分布 (第二章第 3 节表 3).

6. 在上题中的条件下, 证明

$$\lim_{n\to\infty}\mathbf{P}_\theta\left\{\sqrt{\sum_{k=1}^n\xi_{k-1}^2}\,(\widehat{\theta}_n-\theta)\leqslant x\right\}=\begin{cases}\Phi(x), & |\theta|\neq 1,\\ H_\theta^{(2)}(x), & |\theta|=1,\end{cases}$$

其中, $H_\theta^{(2)}(x)$ 是如下的随机变量的分布:

$$\theta\cdot\frac{B_1^2-1}{2\sqrt{\int_0^1 B_s^2ds}}.$$

(这就表明, 如果用以正则化随机变量 $\widehat{\theta}_n-\theta$ 的不是费希尔信息, 而是随机变量 $\left(\sum_{k=1}^n\xi_{k-1}^2\right)^{1/2}$, 那么所得到的极限分布就不是三种, 而是只有两种.)

7. 证明, (第 3 题中的) 极大似然估计 $\widehat{\theta}_n$ 具有如下所示的均值一致渐近性:

$$\sup_\theta\mathbf{E}_\theta|\widehat{\theta}_n-\theta|\to 0,\quad n\to\infty.$$

§5. 沃尔德分解

1. 证明, 具有离散谱 (例如, 谱函数 $F(\lambda)$ 为逐段常数) 的平稳序列是奇异的.

提示: 所考察的序列 $\xi=(\xi_n)$, 其中 $n=0,\pm1,\pm2,\cdots$, 可以表示为如下形式:

$$\xi_n=\sum_{k=-\infty}^{\infty}z_ke^{i\lambda_kn},$$

其中, $z_k=Z(\{\lambda_k\})$, $k=0,\pm1,\pm2,\cdots$ 是相互正交的随机变量, 有 $\mathbf{E}z_k=0$, $\mathbf{E}|z_k|^2=\sigma_k^2$ (采用这些记号, 有 $F(\lambda)=\sum_{k:\,\lambda_k\leqslant\lambda}\sigma_k^2$, $\sum_{k=-\infty}^{\infty}\sigma_k^2<\infty$). 应当证明 $H(\xi)=S(\xi)$, 其中, $H(\xi)$ 是 $L^2(\xi)$ 中的由随机变量 $\xi=(\cdots,\xi_{n-1},\xi_n,\cdots)$ 所生成的闭线性流形, $S(\xi)=\bigcap_{n=-\infty}^{\infty}H_n(\xi)$, 而 $H_n(\xi)$ 是由随机变量 $\xi^n=(\cdots,\xi_{n-1},\xi_n)$ 所生成的.

为证等式 $H(\xi)=S(\xi)$, 只需证明, 对每个 $n\in\mathbb{Z}$, 都有 $\xi_n\in S(\xi)$. 而由平稳性知, 又只需证明 $\xi_0\in S(\xi)$, 亦即证明, 对任何整数 N 与 $\delta>0$, 都存在 $\eta\in H_N(\xi)$, 使得 $\|\xi_0-\eta\|_{H^2}<\delta$.

为此目的, 应当验证, 作为该随机变量, 可以对适当的 $n\leqslant N$, 将其取为 ξ_n. (为此, 对于给定的 $\delta>0$, 应当取 M, 使得 $\sum_{|k|>M}\sigma_k^2<\delta/2$, 再证明

$$\|\xi_n-\xi_0\|_{H^2}\leqslant\frac{2}{3}\delta+\sum_{|k|\leqslant M}\sigma_k^2|e^{i\lambda_kn}-1|,$$

然后确认, 对任何 $N\in\mathbb{Z}$ 和 $\varepsilon>0$, 都存在 $n\leqslant N$, 使得 $|e^{i\lambda_kn}-1|<\varepsilon$, $|k|\leqslant M$.)

2. 设 $\sigma_n^2=\mathbf{E}|\xi_n-\widehat{\xi}_n|^2$, $\widehat{\xi}_n=\widehat{\mathbf{E}}(\xi_n\,|\,H_0(\xi))$. 证明, 如果对某个 $n\geqslant1$, 有 $\sigma_n^2=0$, 则序列 ξ 是奇异的; 而如果有 $\sigma_n^2\to R(0)$, $n\to\infty$, 则序列 ξ 是正则的.

3. 证明, 平稳序列 $\xi=(\xi_n)$, $\xi_n=e^{in\varphi}$ 是奇异的, 其中随机变量 φ 服从区间 $[0,2\pi]$ 上的均匀分布. 试求变量 ξ_n 的线性估计量 $\widehat{\xi}_n$, 并证明非线性估计

$$\widetilde{\xi}_n=\left(\frac{\xi_0}{\xi_{-1}}\right)^n$$

给出了根据 "过去" 样本 $\xi^0=(\cdots,\xi_{-1},\xi_0)$ 对 ξ_n 所作的无误预测, 亦即证明

$$\mathbf{E}|\widetilde{\xi}_n-\xi_n|^2=0,\quad n\geqslant1.$$

提示: 为证序列 $\xi=(\xi_n)$ 的奇异性, 需要证明 $\varepsilon_n=\xi_n/\sqrt{2\pi}$ 是白噪声, 亦即 $\xi_n=\sqrt{2\pi}\,\varepsilon_n$ 具有如 (3) 的表达式.

4. 证明, 关于分解为奇异部分与正则部分的分解式 (1) 是唯一的.

§6. 外推、内插和过滤

1. 证明, 定理 1 中的断言, 在没有诸如 "$\Phi(z)$ 具有收敛半径 $r>1$, $\Phi(z)$ 的零点都分布在区域 $|z|>1$ 中" 之类的假设下, 仍然成立.

2. 证明, 对于正则过程, (4) 式中的函数 $\Phi(z)$ 可以表示为如下形式:

$$\Phi(z)=\sqrt{2\pi}\exp\left\{\frac{1}{2}c_0+\sum_{k=1}^{\infty}c_kz^k\right\},\quad |z|<1,$$

其中

$$c_k=\frac{1}{2\pi}\int_{-\pi}^{\pi}e^{ik\lambda}\ln f(\lambda)d\lambda.$$

并由此推出, 一步上的预报错误 $\sigma_1^2=\mathbf{E}|\widehat{\xi}_1-\xi_1|^2$ 满足如下的塞格 (Szegö) – 柯尔莫戈洛夫公式:

$$\sigma_1^2=2\pi\exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi}\ln f(\lambda)d\lambda\right\}.$$

提示: 关于塞格–柯尔莫戈洛夫公式的证明可按下述步骤进行:

(i) 首先证明

$$\sigma_1^2=\|\widehat{\xi}_1-\xi_1\|_{L^2}^2=\int_{-\infty}^{\infty}|e^{i\lambda}-\widehat{\varphi}_1(\lambda)|^2f(\lambda)d\lambda,\qquad (*)$$

其中, $\widehat{\varphi}_1(\lambda)$ 由 (7) 式给出. 综合 (*) 式, 定理 1 中的符号, 以及 $f(\lambda)=\frac{1}{2\pi}|\Phi(e^{-i\lambda})|^2$, 即知 $\sigma_1^2=|b_0|^2$.

(ii) 由本题的前一部分, 知

$$\Phi(z)=\sum_{k=0}^{\infty}b_kz^k=\sqrt{2\pi}\exp\left\{\frac{1}{2}c_0+\sum_{k=1}^{\infty}c_kz^k\right\},$$

故知 $b_0=\sqrt{2\pi}\exp\left\{\frac{1}{2}c_0\right\}$. 因而

$$\sigma_1^2=2\pi\exp\left\{\frac{1}{2}c_0\right\}=2\pi\exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi}\ln f(\lambda)d\lambda\right\}.$$

3. 在没有假设条件 (22) 的情况下, 证明定理 2.

4. 设不相关的信号 θ 与噪声 η 分别具有谱密度

$$f_\theta(\lambda)=\frac{1}{2\pi}\cdot\frac{1}{|1+b_1e^{-i\lambda}|^2},\quad f_\eta(\lambda)=\frac{1}{2\pi}\cdot\frac{1}{|1+b_2e^{-i\lambda}|^2}.$$

试利用定理 3, 根据值 ξ_k, $k\leqslant n$, 找出变量 θ_{n+m} 的估计 $\widetilde{\theta}_{n+m}$, 其中 $\xi_k=\theta_k+\eta_k$.

将谱密度换为

$$f_\theta(\lambda)=\frac{1}{2\pi}|2+e^{-i\lambda}|^2,\quad f_\eta(\lambda)=\frac{1}{2\pi},$$

再解答同样的问题.

§7. 卡尔曼–布西滤子及其推广

1. 证明, 在情况 (1) 之下, 向量 m_n 与 θ_n-m_n 不相关:

$$\mathbf{E}[m_n^*(\theta_n-m_n)]=0.$$

2. 证明, 在情况 (1) 与 (2) 之下, 量 γ_0 和所有的系数, 至多除了系数 $a_0(n,\xi)$, $A_0(n,\xi)$ 之外, 都与 "随机性" 无关 (即与 ξ 无关). 证明, 此时条件方差 γ_n 亦与 "随机性" 无关: $\gamma_n=\mathbf{E}\gamma_n$.

3. 证明, 方程组 (22) 的解由公式 (23) 所给出.

4. 设 $(\theta,\xi)=(\theta_n,\xi_n)$ 为高斯序列, 满足情况 (1) 中的部分条件:

$$\theta_{n+1}=a\theta_n+b\varepsilon_1(n+1),\quad \xi_{n+1}=A\theta_n+B\varepsilon_2(n+1).$$

证明, 如果 $A\neq 0$, $b\neq 0$, $B\neq 0$, 则滤子的极限误差 $\gamma=\lim\limits_{n\to\infty}\gamma_n$ 存在, 并且等于如下方程的正根:

$$\gamma^2+\left[\frac{B^2(1-a^2)}{A^2}-b^2\right]\gamma-\frac{b^2B^2}{A^2}=0.$$

提示: 如果利用 (8) 式, 则可发现

$$\gamma_{n+1}=b^2+a^2\gamma_n-\frac{a^2\gamma_n^2}{c^2+\gamma_n},$$

其中 $c^2=(\frac{B}{A})^2$. 换言之, 我们有 $\gamma_{n+1}=f(\gamma_n)$, 其中 $f(x)=b^2+a^2x-\frac{a^2x^2}{c^2+x}$, $x\geqslant 0$. 该函数是非降有界的. 有鉴于此, 即可证得, 极限 $\lim\gamma_n$ $(=\gamma)$ 存在, 并且满足方程

$$\gamma^2+[c^2(1-a^2)-b^2]\gamma-b^2c^2=0,$$

由韦达定理知, 该方程有唯一的正根.

5. (内插法; [71, 13.3]) 设 (θ,ξ) 为部分被观察的序列, 遵守正则性关系式 (1) 和 (2).

假设向量 θ_m 的条件分布

$$\pi_a(m,m)=\mathbf{P}(\theta_m\leqslant a\,|\,\mathscr{F}_m^{\xi})$$

为正态的.

(a) 证明, 对于 $n\geqslant m$, 条件分布

$$\pi_a(m,n)=\mathbf{P}(\theta_m\leqslant a\,|\,\mathscr{F}_n^{\xi})$$

仍然是正态的, $\pi_a(m,n)\sim\mathcal{N}(\mu(m,n),\ \gamma(m,n))$.

(b) 试 (根据 $\mathscr{F}_n^{\xi}$) 求 (θ_m 的) 内插估计 $\mu(m,n)$ 和矩阵 $\gamma(m,n)$.

6. (外推法; [71, 13.4]) 设在关系式 (1) 和 (2) 中, 有

$$a_0(n,\xi)=a_0(n)+a_2(n)\xi_n,\quad a_1(n,\xi)=a_1(n),$$
$$A_0(n,\xi)=A_0(n)+A_2(n)\xi_n,\quad A_1(n,\xi)=A_1(n).$$

(a) 证明, 此时条件分布 $\pi_{a,b}(m,m)=\mathbf{P}(\theta_m\leqslant a,\ \xi_n\leqslant b\,|\,\mathscr{F}_m^{\xi}),\ n\geqslant m$, 是正态的.

(b) 试求外推估计

$$\mathbf{E}(\theta_n\,|\,\mathscr{F}_m^{\xi}),\quad 和\quad \mathbf{E}(\xi_n\,|\,\mathscr{F}_m^{\xi}),\quad n\geqslant m.$$

7. (最优控制; [71, 14.3]) 考察 "受控的" 部分被观察的系统 $(\theta_n,\xi_n)_{0\leqslant n\leqslant N}$, 其中

$$\theta_{n+1}=u_n+\theta_n+b\varepsilon_1(n+1),$$
$$\xi_{n+1}=\theta_n+\varepsilon_2(n+1).$$

此处, "控制" u_n 是 $\mathscr{F}_n^{\xi}$ 可测的, 且对一切 $0\leqslant n\leqslant N-1$, 都有 $\mathbf{E}u_n^2<\infty$; 量 $\varepsilon_1(n)$ 与 $\varepsilon_2(n)$ 均与 (1) 和 (2) 中的相同, $n=1,\cdots,N$; $\xi_0=0,\ \theta_0\sim\mathcal{N}(m,\gamma)$.

我们说, "控制" $u^*=(u_0^*,\cdots,u_{N-1}^*)$ 是最优的, 如果 $V(u^*)=\sup\limits_u V(u)$, 其中

$$V(u)=\mathbf{E}\left[\sum_{n=0}^{N-1}(\theta_n^2+u_n^2)+\theta_N^2\right].$$

证明, 最优控制存在, 并且满足下式:

$$u_n^*=-[1+P_{n+1}]^{\oplus}P_{n+1}m_n^*,\quad n=0,1,\cdots,N-1,$$

其中

$$a^{\oplus}=\begin{cases}a^{-1}, & 如果\ a\neq 0,\\ 0, & 如果\ a=0,\end{cases}$$

量 $(P_n)_{0\leqslant n\leqslant N}$ 满足递推关系式

$$P_n=1+P_{n+1}-P_{n+1}^2[1+P_{n+1}]^{\oplus},\quad P_N=1,$$

而 (m_n^*) 由如下的关系式所定义:

$$m_{n+1}^*=u_n^*+\gamma_n^*(1+\gamma_n^*)^{\oplus}(\xi_{n+1}-m_n^*),\quad 0\leqslant n\leqslant N-1,$$

其中 $m_0^*=m$,

$$\gamma_{n+1}^*=\gamma_n^*+1-(\gamma_n^*)^2(1+\gamma_n^*)^{\oplus},\quad 0\leqslant n\leqslant N-1,$$

且 $\gamma_0^*=0$.

8. (变点问题中的非线性滤子; [129]) 在统计检验中, 例如, 在食品质量问题中, 属于观察变量的概率特性问题可以在某个随机时刻 θ 发生改变, 这个时刻就是 (在生产过程中) 发生变化的时刻, 也称为变点. 下面介绍关于发现变点的贝叶斯 (Bayes) 提法, 我们假定所考虑的问题与充分统计量相关联.

设在某个可测空间 $(\Omega,\mathscr{F})$ 中给定了一族概率测度 $\{\mathbf{P}^{\pi}:\ \pi\in[0,1]\}$, 随机变量 θ 的取值集合为 $\{0,1,2,\cdots\}$, 而随机变量观察值序列 $X_1,X_2,\cdots$, 满足如下条件:

(i) $\mathbf{P}^{\pi}\{\theta=0\}=\pi,\ \mathbf{P}^{\pi}\{\theta=k\}=(1-\pi)p_k$, 其中 $p_k\geqslant 0,\ \sum\limits_{k=1}^{\infty}p_k=1$.

(ii) 对于每个 $\pi\in[0,1]$ 和 $n\geqslant 1$, 都有

$$\begin{aligned}\mathbf{P}^{\pi}\{X_1\leqslant x_1,\cdots,X_n\leqslant x_n\}&=\pi\mathbf{P}^1\{X_1\leqslant x_1,\cdots,X_n\leqslant x_n\}\\&+(1-\pi)\sum_{k=0}^{n-1}p_{k+1}\mathbf{P}^0\{X_1\leqslant x_1,\cdots,X_k\leqslant x_k\}\mathbf{P}^1\{X_{k+1}\leqslant x_{k+1},\cdots,X_n\leqslant x_n\}\\&+(1-\pi)(p_{n+1}+p_{n+2}+\ldots)\mathbf{P}^0\{X_1\leqslant x_1,\cdots,X_n\leqslant x_n\},\quad x_k\in\mathbb{R}.\end{aligned}$$

(iii) $\mathbf{P}^j\{X_1\leqslant x_1,\cdots,X_n\leqslant x_n\}=\prod\limits_{k=1}^{n}\mathbf{P}^j\{X_k\leqslant x_k\},\ j=0,1.$

(这些假设 (i) ~ (iii) 的直观意义是: 如果 $\theta=0$ 或 $\theta=1$, 则从一开始观察, 变点就已经存在. 此时与生产过程中的变点相应的量 $X_1,X_2,\cdots$ 是独立同分布的, 均服从分布函数 $F_1(x)=\mathbf{P}^1\{X_1\leqslant x\}$. 如果 $\theta>n$, 则变点在时刻 n 以后出现, 从而 $X_1,\cdots,X_n$ 对应于生产中的正常步骤, 是独立同分布的随机变量序列, 均服从分布函数 $F_0(x)=\mathbf{P}^0\{X_1\leqslant x\}$. 如果 $\theta=k$, 其中 $1<k\leqslant n$, 则随机变量 $X_1,\cdots,X_{k-1}$ 独立同分布, 服从分布函数 $F_0(x)$, 而随机变量 $X_k,\cdots,X_n$ 独立同分布, 但分布函数却是 $F_1(x)$. 我们假设 $F_0(x)\neq F_1(x)$.)

设 $f_0(x)$ 与 $f_1(x)$ 分别是 $F_0(x)$ 与 $F_1(x)$ 分布密度 (关于某个被 $F_0(x)$ 与 $F_1(x)$ 所控制的分布, 例如, $(F_0(x)+F_1(x))/2$).

以 τ 表示 (关于 $(\mathscr{F}_n)_{n\geqslant 0}$, 其中 $\mathscr{F}_0=\{\varnothing,\Omega\}$, $\mathscr{F}_n=\sigma(X_1,\cdots,X_n)$ 的) 马尔可夫时, 它是关于变点发生的报警信号的发出时刻. 作为这样的时刻, 可以用两个变量来刻画其特性: 其一是假警报概率 $\mathbf{P}^{\pi}\{\tau<\theta\}$, 其二是在警报信号正确发出的情况下 (意即 $\tau>\theta$), 与变点发生时刻的平均延迟时间 $\mathbf{E}^{\pi}(\tau-\theta)^+$.

自然希望能找到这样的时刻 τ, 它能够同时减小假警报概率和平均延迟时间. 但由于这样的时刻不存在 (平凡情况除外), 人们引入贝叶斯风险 $(c>0)$:

$$R^{\pi}(\tau)=\mathbf{P}^{\pi}\{\tau<\theta\}+c\,\mathbf{E}^{\pi}(\tau-\theta)^+.$$

称时刻 τ^* 是最佳的, 如果对于一切 $\pi\in[0,1]$, 都有 $\mathbf{P}^{\pi}\{\tau^*<\infty\}=1$, 并且对于所有 $\mathbf{P}^{\pi}$ 有限的马尔可夫时 τ, 都有不等式 $R^{\pi}(\tau^*)\leqslant R^{\pi}(\tau)$ 成立.

根据第八章第 9 节第 8 题, 这样的时刻 τ^* 是存在的, 并且 (在 $p_k=(1-$

$p)^{k-1}p$, $0 < p \leqslant 1$, $k \geqslant 1$ 时) 具有如下的简单结构:

$$\tau^* = \inf\{n \geqslant 0:\ \pi_n \geqslant A\},$$

其中, A 是某个与 c 和 p 有关的阈值常数, π_n 是在时刻 n 以前出现变点的后验概率:

$$\pi_n = \mathbf{P}^\pi\{\theta \leqslant n \,|\, \mathscr{F}_n\}, \quad n \geqslant 0, \quad \pi_0 = \pi.$$

(a) 证明, 后验概率 π_n, $n \geqslant 0$, 满足如下的递推关系式:

$$\pi_{n+1} = \frac{\pi_n f_1(X_{n+1}) + p(1-\pi_n) f_1(X_{n+1})}{\pi_n f_1(X_{n+1}) + p(1-\pi_n) f_1(X_{n+1}) + (1-\pi)(1-\pi_n) f_0(X_{n+1})}.$$

(b) 证明, 如果 $\varphi_n = \frac{\pi_n}{1-\pi_n}$, 则

$$\varphi_{n+1} = (p + \varphi_n)\frac{f_1(X_{n+1})}{(1-p) f_0(X_{n+1})}, \quad \varphi_0 = \frac{\pi}{1-\pi}.$$

(c) 记 $\varphi = \varphi_n(p)$, 并令 $\pi = 0$ 和 $\psi = \lim\limits_{p \downarrow 0} \frac{\varphi_n(p)}{p}$, 证明

$$\psi_{n+1} = (1 + \psi_n)\frac{f_1(X_{n+1})}{f_0(X_{n+1})}, \quad \psi_0 = 0.$$

注: 如果令 $\theta_n = I(\theta \leqslant n)$, 则 $\pi_n = \mathbf{E}^\pi(\theta_n \,|\, \mathscr{F}_n)$, 是基于观察值 $X_1, \cdots, X_n$ 的 θ_n 的均方意义下的最优估计. 由 (a)(b) 和 (c), 我们看出, 统计量 π_n, φ_n 和 ψ_n 满足非线性递推关系式, 按照通常的说法, (在根据观察值 $X_1, \cdots, X_n$ 估计过程 $(\theta_n)_{n\geqslant 0}$ 中的值 θ 的问题中), 它们构成了非线性滤子.

(d) 证明, 序列 $(\pi_n)_{n\geqslant 0}$, $(\varphi_n)_{n\geqslant 0}$ 与序列 $(\psi_n)_{n\geqslant 0}$ 都分别 (关于 $(\mathscr{F})_{n\geqslant 0}$ 与概率测度 $\mathbf{P}^\pi$, $\pi \in [0,1]$) 是马尔可夫链, 而 $(\psi_n)_{n\geqslant 0}$(关于 $(\mathscr{F})_{n\geqslant 0}$ 与测度 $\mathbf{P}^0$) 是马尔可夫链.

第七章　构成鞅的随机变量序列

§1. 鞅和相关概念的定义

1. 试证明条件 (2) 与条件 (3) 的等价性.

提示: 应当采用反证法.

2. 设 σ 与 τ 都是马尔可夫时, 证明 $\tau + \sigma$, $\tau \vee \sigma$, $\tau \wedge \sigma$ 也都是马尔可夫时; 并且如果对一切 $\omega \in \Omega$, 都有 $\sigma(\omega) \leqslant \tau(\omega)$, 则有 $\mathscr{F}_\sigma \subseteq \mathscr{F}_\tau$. 试问, 如果仅有概率为 1 地成立 $\sigma \leqslant \tau$, 那么这一性质能否保持?

提示: 如果对一切 $\omega \in \Omega$, 都有 $\sigma(\omega) \leqslant \tau(\omega)$, 则对任何 $A \in \mathscr{F}_\sigma$, 都有

$$A \cap \{\tau = n\} = A \cap \{\sigma \leqslant n\} \cap \{\tau = n\} \in \mathscr{F}_n,$$

从而 $A \in \mathscr{F}_\tau$.

3. 证明, τ 与 X_τ 都是 $\mathscr{F}_\tau$ 可测的.

4. 设 $Y = (Y_n,\ \mathscr{F}_n)$ 是鞅 (半鞅), 而 $V = (V_n,\ \mathscr{F}_{n-1})$ 是可预报序列, 且 $(V \cdot Y)_n$ 是可积随机变量, $n \geqslant 0$. 证明, $V \cdot Y$ 是鞅 (半鞅).

5. 设 $\mathscr{G}_1 \supseteq \mathscr{G}_2 \supseteq \cdots$ 是非升的 σ-代数族, ξ 是可积的随机变量. 令 $X_n = \mathbf{E}(\xi \,|\, \mathscr{G}_n)$. 证明, 序列 $(X_n)_{n\geqslant 1}$ 形成逆鞅, 即对任何 $n \geqslant 1$, 都有

$$\mathbf{E}(X_n \,|\, X_{n+1}, X_{n+2}, \cdots) = X_{n+1} \quad (\mathbf{P}\text{-a.s.}).$$

6. 设 $\xi_1, \xi_2, \cdots$ 为独立随机变量序列, 有 $\mathbf{P}\{\xi_i = 0\} = \mathbf{P}\{\xi_i = 2\} = \frac{1}{2}$. 令 $X_n = \prod\limits_{i=1}^{n} \xi_i$. 证明, 不存在可积的随机变量 ξ 和非降的 σ-代数族 $(\mathscr{F}_n)_{n\geqslant 1}$, 使得 $X_n =$

$\mathbf{E}(\xi\,|\,\mathscr{F}_n)$ ($\mathbf{P}$-a.s.), $n \geqslant 1$. (这个例子说明, 并非所有的鞅序列 $(X_n)_{n\geqslant 1}$ 都可以表示成 $(\mathbf{E}(\xi\,|\,\mathscr{F}_n))_{n\geqslant 1}$ 的形式; 试比较第一章第 11 节中的例 3.)

提示: 应当采用反证法.

7. (a) 设 $\xi_1, \xi_2, \cdots$ 为独立随机变量序列, 有 $\mathbf{E}|\xi_n| < \infty$, $\mathbf{E}\xi_n = 0$, $n \geqslant 1$. 证明, 对每个 k, 如下的序列都形成鞅:

$$X_n^{(k)} = \sum_{1\leqslant i_1<\cdots<i_k\leqslant n} \xi_{i_1}\cdots\xi_{i_k}, \quad n \geqslant k.$$

(b) 设 $\xi_1, \xi_2, \cdots$ 为可积随机变量序列, 使得

$$\mathbf{E}(\xi_{n+1}\,|\,\xi_1,\cdots,\xi_n) = \frac{\xi_1+\cdots+\xi_n}{n} \quad (= X_n).$$

证明, 序列 $X_1, X_2, \cdots$ 形成鞅.

8. 试举例说明, 存在这样的鞅序列 $X = (X_n,\ \mathscr{F}_n)_{n\geqslant 1}$, 其中的随机变量序列 $\{X_n,\ n \geqslant 1\}$ 不是一致可积的.

9. 设 $X = \{X_n\}_{n\geqslant 0}$ 是具有可数状态集合 $E = \{i, j, \cdots\}$ 与转移概率 p_{ij} 的马尔可夫链 (第八章第 1 节). 令 $\psi = \psi(x)$, $x \in E$ 是有界函数, 使得对于 $\lambda > 0$ 和 $i \in E$, 有 $\sum_{j\in E} p_{ij}\psi(j) \leqslant \lambda\psi(i)$ (这样的函数称为 λ-峰态的, 或 λ-调和的). 证明, 序列 $\{\lambda^{-n}\psi(X_n)\}_{n\geqslant 0}$ 是上鞅.

10. 设 $\tau_1, \tau_2, \cdots$ 为停时序列, 使得逐点地有 $\tau_n \downarrow \tau$ 或逐点地有 $\tau_n \uparrow \tau$, 则在两种情况下, τ 都是停时.

11. 若 σ 与 τ 都是停时, 则有

$$\mathscr{F}_{\sigma\wedge\tau} = \mathscr{F}_\sigma \cap \mathscr{F}_\tau, \quad \mathscr{F}_{\sigma\vee\tau} = \sigma(\mathscr{F}_\sigma \cup \mathscr{F}_\tau).$$

12. 设 σ 为 (有限) 停时, 而 $\tau_1, \tau_2, \cdots$ 为停时序列, 有 $\tau_n \uparrow \infty$. 证明,

$$\mathscr{F}_{\sigma\wedge\tau_n} \uparrow \mathscr{F}_\sigma.$$

而若 $\tau_n \downarrow \tau$, 则有 $\mathscr{F}_\tau = \bigcap_n \mathscr{F}_{\tau_n}$.

13. 设 $\xi_1, \xi_2, \cdots$ 为相互独立的标准高斯随机变量序列 ($\xi_n \sim \mathscr{N}(0,1)$), 而 $S_n = \xi_1 + \cdots + \xi_n$, $n \geqslant 1$. 令

$$X_n = \frac{1}{\sqrt{n+1}} \exp\left\{\frac{S_n^2}{2(n+1)}\right\},$$

$\mathscr{F}_n^\xi = \sigma(\xi_1, \cdots, \xi_n)$. 证明, 序列 $(X_n, \mathscr{F}_n^\xi)_{n\geqslant 1}$ 是鞅.

14. 设 $X = (X_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 为随机序列, $\Delta X_n = X_n - X_{n-1}$, $n \geqslant 1$, $\nu(\omega;\ \{n\}\times dx) = \mathbf{P}(\Delta X_n \in dx\,|\,\mathscr{F}_{n-1})(\omega)$ 是条件期望的正则支. 对 $u \in \mathbb{R}$, 令

$$A(u)_0 = 0, \quad A(u)_n = \sum_{1\leqslant k\leqslant n} \int (e^{iux} - 1)\nu(\omega;\ \{k\}\times dx),$$

$$M_n(u) = e^{iuX_n} - \sum_{k=1}^{n} e^{iuX_{k-1}} \Delta A(u)_k.$$

证明, 过程 $M(u) = (M_n(u),\ \mathscr{F}_n)_{n\geqslant 1}$ 是鞅.

15. 符号同上题, 对 $u \in \mathbb{R}$, 令

$$G(u)_0 = 0, \quad G(u)_n = \prod_{1\leqslant k\leqslant n} \int e^{iux}\nu(\omega;\ \{k\}\times dx), \quad n \geqslant 1.$$

假设 $G(u)_n > 0$, $n \geqslant 1$. 证明, 序列

$$\left(\frac{e^{iuX_n}}{G(u)_n}, \mathscr{F}_n\right)_{n\geqslant 1}, \quad \text{即序列} \quad \left(\frac{e^{iuX_n}}{\prod_{k=1}^{n} \mathbf{E}(e^{iu\Delta X_k}\,|\,\mathscr{F}_{k-1})}, \mathscr{F}_n\right)_{n\geqslant 1}$$

是鞅.

16. 设 $X = (X_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 为随机序列, 对某个 $c > 0$ 和一切 $n \geqslant 1$, 都有 $|\Delta X_n| \leqslant c$ ($\mathbf{P}$-a.s.). 令

$$Y_n = \frac{e^{X_n}}{\prod_{i=1}^{n} \mathbf{E}(e^{\Delta X_i}\,|\,\mathscr{F}_{i-1})},$$

其中 $\Delta X_i = X_i - X_{i-1}$, $i \geqslant 1$. 证明, (实值) 序列 $Y = (Y_n,\ \mathscr{F}_n)_{n\geqslant 1}$ 是鞅 (试比较上面的第 15 题).

如果取消关于 $|\Delta X_n|$ 的有界性条件, 那么结论是否仍然成立?

17. 设 $S_0 = 0$, $S_k = \xi_1 + \cdots + \xi_k$, $1 \leqslant k \leqslant n$, 其中 $\xi_1, \cdots, \xi_n$ 是相互独立的标准正态 $\mathscr{N}(0,1)$ 随机变量. 记 $\Phi(x) = \mathbf{P}\{\xi_1 \leqslant x\}$, $\mathscr{F}_k = \sigma(\xi_1, \cdots, \xi_k)$, $1 \leqslant k \leqslant n$, $\mathscr{F}_0 = \{\varnothing,\ \Omega\}$. 对 $a \in \mathbb{R}$, 令

$$X_k = \Phi\left(\frac{a - S_k}{\sqrt{n-k}}\right).$$

证明, 对任何 $a \in \mathbb{R}$, 序列 $X = (X_k,\ \mathscr{F}_k)_{0\leqslant k\leqslant n}$ 都是鞅.

18. 设 $S_0 = 0$, $S_k = \xi_1 + \cdots + \xi_k$, $1 \leqslant k \leqslant n$, 其中 $\xi_1, \cdots, \xi_n$ 是独立同分布的随机变量, 服从某个对称的分布. 记 $F(x;\ k) = \mathbf{P}\{S_k \leqslant x\}$. 作为上题的推广, 对 $a \in \mathbb{R}$, 令 $X_k = F(a - S_k,\ n - k)$. 证明, 序列 $X = (X_k,\ \mathscr{F}_k)_{0\leqslant k\leqslant n}$ 仍是鞅. (该结论的一个应用, 参阅第 2 节第 12 题.)

19. 设 $\xi_1, \xi_2, \cdots$ 是独立同分布的随机变量序列, 服从分布 $F = F(x)$, $x \in \mathbb{R}$. 设

$$F_n(x;\ \omega) = \frac{1}{n}\sum_{k=1}^{n} I(\xi_k(\omega) \leqslant x), \quad x \in \mathbb{R}$$

为经验分布函数, $n \geqslant 1$ (参阅第三章第 13 节). 对 $x \in \mathbb{R}$, 令

$$Y_n(x) = F_n(x;\ \omega) - F(x), \quad \mathscr{G}_n(x) = \sigma\,(Y_n(x), Y_{n+1}(x), \cdots).$$

证明, 对任何 $x \in \mathbb{R}$, 序列 $(Y_n(x),\ \mathscr{G}_n(x))_{n\geqslant 1}$ 都是逆鞅.

20. 设 $X = (X_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 与 $Y = (Y_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 为两个半鞅.

(a) 证明, 序列 $X \vee Y = (X_n \vee Y_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 仍然为半鞅.

(b) 试问, 序列

$$X + Y = (X_n + Y_n,\ \mathscr{F}_n)_{n\geqslant 0} \quad 与 \quad XY = (X_n Y_n,\ \mathscr{F}_n)_{n\geqslant 0}$$

是否为半鞅? (如果 "是", 那么在怎样的条件之下? 如果 "否", 那么为什么?)

(c) 如果 X 与 Y 是两个鞅, 或者 X 与 Y 是两个上鞅, 再讨论上述问题.

21. 设 $\xi_1, \xi_2, \cdots$ 是可交换随机变量的无限序列(即对任何 $n \geqslant 1$, 随机向量 $(\xi_1, \cdots, \xi_n)$ 都与随机向量 $(\xi_{\pi_1}, \cdots, \xi_{\pi_n})$ 同分布, 其中 $(\pi_1, \cdots, \pi_n)$ 是 $(1, \cdots, n)$ 的任意一种排列). 令 $\mathbf{E}\xi_1 < \infty$, $S_n = \xi_1 + \cdots + \xi_n$, $\mathscr{G}_n = \sigma\left(\frac{S_n}{n}, \frac{S_{n+1}}{n+1}, \cdots\right)$.

作为第一章第 11 节例 4 中所讨论的内容的推广, 证明,

$$\mathbf{E}\left(\frac{S_n}{n}\,\middle|\,\mathscr{G}_{n+1}\right) = \frac{S_{n+1}}{n+1} \quad (\mathbf{P}\text{-a.s.}), \quad n \geqslant 1,$$

即序列 $\left(\frac{S_n}{n},\ \mathscr{G}_n\right)_{n\geqslant 1}$ 是逆鞅.

22. 证明, 一切逆鞅都是一致可积的.

23. 根据定义, 马尔可夫时 τ 的 σ-代数 $\mathscr{F}_\tau$ 是集合类 $\{A \in \mathscr{F}:\ A \cap \{\tau = n\} \in \mathscr{F}_n,\ 一切\ n \geqslant 0\}$. 试问: 为何不将其定义为 $\mathscr{F}_\tau \overset{\text{def}}{=} \sigma\,(\mathscr{F}_n : n \leqslant \tau)$?

24. 如果序列 $X = (X_n,\ \mathscr{F}_n)_{n\geqslant 1}$ 是鞅, 而 $(n_k) \subseteq (n)$ 是某个子列, 则 $(X_{n_k},\ \mathscr{F}_{n_k})_{k\geqslant 1}$ 仍然是鞅. 试举例说明, 一般来说, 对于局部鞅该结论未必成立.

25. 在鞅论中, 如果 $\Pi = (\Pi_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 是一致可积的非负上鞅, 并且逐点地有 $\Pi_n(\omega) \to 0$, $n \to \infty$, $\omega \in \Omega$, 则称其为位势.

证明, 对于每个位势, 都唯一存在可预报的非降序列 $A = (A_n,\ \mathscr{F}_{n-1})_{n\geqslant 0}$, 使得

$$\Pi_n = \mathbf{E}\,(A_\infty - A_n \mid \mathscr{F}_n), \quad n \geqslant 0,$$

其中 $A_0 = 0$, $\mathscr{F}_{-1} = \mathscr{F}_0$.

26. 设 $X = (X_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 是上鞅. 证明, 如下各条件相互等价:

(i) 存在半鞅序列 $Y = (Y_n,\ \mathscr{F}_n)_{n\geqslant 0}$, 使得对每个 $n \geqslant 0$, 都有 $X_n \geqslant Y_n$ ($\mathbf{P}$-a.s.).

(ii) 存在唯一的里斯分解:

$$X_n = M_n + \Pi_n,$$

其中 $M = (M_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 是鞅, 而 $\Pi = (\Pi_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 是位势的非负上鞅.

27. 设 $X = (X_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 是半鞅. 证明, 存在非负鞅 $M = (M_n,\ \mathscr{F}_n)_{n\geqslant 0}$, 使得

$$X_n^+ \leqslant M_n, \quad n \geqslant 0 \quad 和 \quad \sup_n \mathbf{E}X_n^+ = \sup_n \mathbf{E}M_n.$$

提示: 利用 $X^+ = (X_n^+,\ \mathscr{F}_n)_{n\geqslant 0}$ 也是半鞅的事实, 并令

$$M_n = \lim_{m\to\infty} \mathbf{E}\,(X_{n+m}^+ \mid \mathscr{F}_n).$$

28. 设 $(\Omega, \mathscr{F}, \mathbf{P})$ 是带有滤子 $(\mathscr{F}_n)_{n\geqslant 0}$ 的概率空间. 设 σ 与 τ 是两个马尔可夫时, 有 $\sigma(\omega) \leqslant \tau(\omega)$, $\omega \in \Omega$. 令 $A = \{\omega:\ \sigma(\omega) < n \leqslant \tau(\omega)\}$. 证明, 对每个 $n \geqslant 1$, 都有 $A \in \mathscr{F}_{n-1}$. 换言之, 序列 $(X_n)_{n\geqslant 1}$ 是可预报的, 其中

$$X_n = \begin{cases} 1, & 如果\ \sigma(\omega) < n \leqslant \tau(\omega), \\ 0, & 其他场合 \end{cases}$$

(即 X_n 是 $\mathscr{F}_{n-1}$ 可测的, $n \geqslant 1$).

29. (推广定理 2) 设 $X = (X_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 是半鞅, 具有杜布分解 $X_n = m_n + A_n$, $n \geqslant 0$, 其中 $A_0 = 0$, $m_0 = X_0$. 证明, 如果随机变量族 $\{X_n,\ n \geqslant 0\}$ 一致可积, 则有 $\mathbf{E}A_\infty < \infty$, 并且随机变量族 $\{m_n,\ n \geqslant 0\}$ 也一致可积.

30. 设 $M = (M_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 是平方可积鞅. 证明,

$$\sup_n \mathbf{E}M_n^2 < \infty \quad \Leftrightarrow \quad \sum_{k\geqslant 1} \mathbf{E}(M_k - M_{k-1})^2 < \infty.$$

31. 设 $\tau = \tau(\omega)$ 是关于流 $(\mathscr{F}_n)_{n\geqslant 0}$ 的马尔可夫时. 设 $f = f(n)$ 是定义在 $n \in \mathbb{N} = \{0, 1, 2, \cdots\}$ 上的非降函数, 有 $f(n) \geqslant n$. 证明, $\widetilde{\tau}(\omega) = f(\tau(\omega))$ 也是马尔可夫时.

32. 设 $X = (X_n,\ \mathscr{F}_n)$ 是关于测度 $\mathbf{P}$ 的鞅, 而测度 $\mathbf{Q}$ 等价于测度 $\mathbf{P}$ ($\mathbf{Q} \sim \mathbf{P}$). 试举例说明, 关于 $\mathbf{Q}$, 序列 $X = (X_n,\ \mathscr{F}_n)$ 可能不是鞅.

33. 根据例 5, 如果 $X = (X_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 是半鞅, 而 $g = g(x)$ 是非降的下凸函数, 使得 $\mathbf{E}|g(X_n)| < \infty$, $n \geqslant 0$, 则序列 $(g(X_n),\ \mathscr{F}_n)_{n\geqslant 0}$ 仍然是半鞅. 试举例说明, 如果函数 $g = g(x)$ 下凸, 但不非降, 则序列 $(g(X_n),\ \mathscr{F}_n)_{n\geqslant 0}$ 不是半鞅.

§2. 在时间变量为随机时间时鞅性的不变性

1. 证明, 如果将定理 1 中的条件 (4) 换为如下条件, 那么定理的结论可对半鞅序列成立:

$$\underline{\lim}_{n\to\infty}\int_{\{\tau_2>n\}} X_n^+ d\mathbf{P}=0.$$

提示: 证明过程本质上与定理 1 的证明相同, 只是需要指出: 既然 $X_m \leqslant X_m^+$, 故有如下的不等式成立:

$$\int_{B\cap\{\tau_2\geqslant n\}} X_{\tau_2} d\mathbf{P} \geqslant \overline{\lim}_{m\to\infty}\left[\int_{B\cap\{\tau_2\geqslant n\}} X_n d\mathbf{P}-\int_{B\cap\{\tau_2>m\}} X_m d\mathbf{P}\right]$$
$$\geqslant \int_{B\cap\{\tau_2\geqslant n\}} X_n d\mathbf{P}-\underline{\lim}_{m\to\infty}\int_{B\cap\{\tau_2>m\}} X_m^+ d\mathbf{P}.$$

2. 设 $X=(X_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 是平方可积鞅, τ 是停时, 且 $\mathbf{E}X_0=0$,

$$\underline{\lim}_{n\to\infty}\int_{\{\tau>n\}} X_n^2 d\mathbf{P}=0.$$

证明,

$$\mathbf{E}X_\tau^2=\mathbf{E}\langle X\rangle_\tau \quad \left(=\mathbf{E}\sum_{j=0}^{\tau}(\Delta X_j)^2\right),$$

其中, $\Delta X_0=X_0,\ \Delta X_j=X_j-X_{j-1},\ j\geqslant 1$.

提示: 为证明不等式

$$\mathbf{E}X_\tau^2\leqslant \mathbf{E}\sum_{j=0}^{\tau}(\Delta X_j)^2,$$

应当利用定理 1 和法图引理 ($\mathbf{E}\,\underline{\lim}_{N\to\infty} X_{\tau\wedge N}^2 \leqslant \underline{\lim}_{N\to\infty} \mathbf{E}X_{\tau\wedge N}^2$). 为证明反过来的不等式, 应当利用

$$\mathbf{E}X_\tau^2\geqslant \mathbf{E}X_{\tau\wedge N}^2=\mathbf{E}\sum_{j=0}^{\tau\wedge N}(\Delta X_j)^2,$$

以及法图引理.

3. 证明, 对任何鞅与任何非负半鞅 $X=(X_n,\ \mathscr{F}_n)_{n\geqslant 0}$, 以及任何停时 τ, 都有

$$\mathbf{E}|X_\tau|\leqslant \lim_{n\to\infty}\mathbf{E}|X_n|.$$

提示: 注意 $|X|$ 是半鞅, 由定理 1 知, 对一切 $N\geqslant 1$, 都有 $\mathbf{E}|X_{\tau\wedge N}|\leqslant \mathbf{E}|X_N|$. 因此 $\underline{\lim}_N \mathbf{E}|X_{\tau\wedge N}|\leqslant \underline{\lim}_N \mathbf{E}|X_N|$, 再利用法图引理.

4. 设 $X=(X_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 是上鞅, 有 $X_n\geqslant \mathbf{E}(\xi\,|\,\mathscr{F}_n)$ ($\mathbf{P}$-a.s.), $n\geqslant 0$, 其中 $\mathbf{E}|\xi|<\infty$. 证明, 如果 τ_1 与 τ_2 是两个停时, 满足条件 $\mathbf{P}\{\tau_1\leqslant\tau_2\}=1$, 则有

$$X_{\tau_1}\geqslant \mathbf{E}(X_{\tau_2}\,|\,\mathscr{F}_{\tau_1})\quad (\mathbf{P}\text{-a.s.}).$$

提示: 利用定理 1 的结果, 验证在我们的题目中, 有 $\mathbf{E}|X_{\tau_1}|<\infty$ 和 $\mathbf{E}|X_{\tau_2}|<\infty$, 以及

$$\underline{\lim}_{n}\int_{\{\tau_2>n\}}|X_n|d\mathbf{P}=0.$$

5. 设 $\xi_1,\xi_2,\cdots$ 为相互独立的随机变量序列, 有 $\mathbf{P}\{\xi_k=1\}=\mathbf{P}\{\xi_k=-1\}=1/2$. 而 a 与 b 为正数, 有 $b>a$. 令

$$X_n=a\sum_{k=1}^{n}I(\xi_k=1)-b\sum_{k=1}^{n}I(\xi_k=-1)$$

和

$$\tau=\inf\{n\geqslant 1:\ X_n\leqslant -r\},\quad r>0.$$

证明, 当 $\lambda\leqslant\alpha_0$ 时, 有 $\mathbf{E}\,e^{\lambda\tau}<\infty$; 当 $\lambda>\alpha_0$ 时, 有 $\mathbf{E}\,e^{\lambda\tau}=\infty$, 其中

$$\alpha_0=\frac{b}{a+b}\ln\frac{2b}{a+b}+\frac{a}{a+b}\ln\frac{2a}{a+b}.$$

6. 设 $\xi_1,\xi_2,\cdots$ 为相互独立的随机变量序列, 有 $\mathbf{E}\xi_j=0$, $\mathbf{D}\,\xi_j=\sigma_j^2$. 令 $S_n=\xi_1+\cdots+\xi_n$, $\mathscr{F}_n^\xi=\sigma(\xi_1,\cdots,\xi_n)$, $n\geqslant 1$. 证明对于沃尔德恒等式 (13) 和 (14) 的如下推广:

$$\text{如果 } \mathbf{E}\sum_{j=1}^{\tau}\mathbf{E}|\xi_j|<\infty,\quad \text{则 } \mathbf{E}\,S_\tau=0;$$

$$\text{如果 } \mathbf{E}\sum_{j=1}^{\tau}\mathbf{E}\xi_j^2<\infty,\quad \text{则 } \mathbf{E}\,S_\tau^2=\mathbf{E}\sum_{j=1}^{\tau}\xi_j^2=\mathbf{E}\sum_{j=1}^{\tau}\sigma_j^2.$$

7. 设 $X=(X_n,\ \mathscr{F}_n)_{n\geqslant 1}$ 是平方可积鞅, τ 是停时. 证明,

$$\mathbf{E}\,X_\tau^2\leqslant \mathbf{E}\sum_{n=1}^{\tau}(\Delta X_n)^2.$$

再证明, 如果

$$\underline{\lim}_{n\to\infty}\mathbf{E}\,(X_n^2 I(\tau>n))<\infty \quad \text{或} \quad \underline{\lim}_{n\to\infty}\mathbf{E}\,(|X_n|I(\tau>n))=0,$$

则 $\mathbf{E}\,(\Delta X_\tau)^2=\mathbf{E}\sum\limits_{n=1}^{\tau}X_n^2$.

8. 设 $X=(X_n,\ \mathscr{F}_n)_{n\geqslant 1}$ 是半鞅, $\tau_1\leqslant\tau_2\leqslant\cdots$ 是停时序列, 使得 $\mathbf{E}\,X_{\tau_m}$ 确定和

$$\lim_{n\to\infty}\mathbf{E}\,(X_n^+I(\tau_m\geqslant n))=0,\quad m\geqslant 1.$$

证明, 序列 $(X_{\tau_m},\ \mathscr{F}_{\tau_m})_{m\geqslant 1}$ 仍然为半鞅 (如同往常, $\mathscr{F}_{\tau_m}=\{A\in\mathscr{F}:\ A\cap\{\tau_m=j\}\in\mathscr{F}_j,\ j\geqslant 1\}$).

9. 设 $X=(X_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 为非负半鞅, $\tau_0\leqslant\tau_1\leqslant\cdots$ 是停时. 证明, 序列 $(X_{\tau_n},\ \mathscr{F}_{\tau_n})_{n\geqslant 1}$ 仍为半鞅①.

10. 作为更新理论的基本定理 (见第 2 节第 4 小节) 的补充, 试证明, 如果 $0<\mathbf{D}\,\sigma_1<\infty$, 则有

$$\frac{\mathbf{D}\,N_t}{t}\to a^3\mathbf{D}\,\sigma_1,\quad t\to\infty$$

和中心极限定理

$$\mathbf{P}\left\{\frac{N_t-at}{\sqrt{a^3\mathbf{D}\,\sigma_1 t}}\leqslant x\right\}\to\Phi(x),\quad t\to\infty$$

成立, 其中 $a=(\mathbf{E}\,\sigma_1)^{-1}$.

11. 设 $\xi_1,\xi_2,\cdots$ 为独立同分布随机变量序列. 令 $S_n=\xi_1+\cdots+\xi_n,\ n\geqslant 1$,

$$\tau=\inf\{n\geqslant 1:\ S_n>0\}$$

(如果对一切 $n\geqslant 1$, 都有 $S_n\leqslant 0$, 则如同往常, 令 $\tau=\infty$).

证明, 如果 $\mathbf{E}\xi_1=0$, 则 $\mathbf{E}\,\tau=\infty$.

12. 试利用第 1 节第 18 题中的序列 $X=(X_k,\ \mathscr{F}_k)_{0\leqslant k\leqslant n}$ 的鞅性质, 和关于停时 $\tau_a=\min\{0\leqslant k\leqslant n:\ S_k>a\},\ a>0$ (如果对一切 $0\leqslant k\leqslant n$, 都有 $S_k\leqslant a$, 则令 $\tau_a=n+1$) 的性质 $\mathbf{E}\,X_0=\mathbf{E}\,X_{\tau_a}$, 证明如下不等式 (参阅第四章第 4 节引理 1),

$$\mathbf{P}\left\{\max_{0\leqslant k\leqslant n}S_k>a\right\}\leqslant 2\mathbf{P}\{S_n>a\}.$$

13. 作为定理 1 和 2 中命题的补充, 试证明如下结论: 设 $X=(X_n,\ \mathscr{F}_n)$ 是鞅, τ 为停时 $(\mathbf{P}\{\tau<\infty\}=a)$, 有 $\mathbf{E}\,|X_\tau|<\infty$, $\lim\limits_{n\to\infty}\mathbf{E}\,[|X_n|I(\tau>n)]=0$. 则有

$$\lim_{n\to\infty}\mathbf{E}\,[|X_\tau|I(\tau>n)]=0,$$
$$\mathbf{E}\,|X_\tau-X_{\tau\wedge n}|\to 0,\quad n\to\infty$$

和 $\mathbf{E}\,X_\tau=\mathbf{E}\,X_0$.

①原书中为 "序列 $(X_{\tau_n},\ \mathscr{F}_n)_{n\geqslant 1}$ 仍然为半鞅" —— 译者注.

14. (续定理 1 和 2) 设 τ_1 与 τ_2 是两个有限停时, 有 $\mathbf{P}\{\tau_1\leqslant\tau_2\}=1$, 而 $X=(X_n)_{n\geqslant 0}$ 是鞅. 证明, 如果

$$\mathbf{E}\sup_{n\leqslant\tau_2}|X_n|<\infty,\qquad (*)$$

则有 $\mathbf{E}\,(X_{\tau_2}\,|\,\mathscr{F}_{\tau_1})=X_{\tau_1}$ ($\mathbf{P}$-a.s.).

提示: 利用如下事实, 在条件 (*) 之下, 随机变量族 $\{|X_{\tau_2\wedge k}|,\ k\geqslant 0\}$ 是一致可积的.

15. 设 τ 是由例 1 中 (16) 式所定义的停时. 证明, 对一切 $p\geqslant 1$, 都有 $\mathbf{E}\,\tau^p<\infty$.

16. 试举例说明, 存在鞅序列 $X=(X_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 和停时 τ, 满足如下条件 (参阅定理 1):

$$\lim_{n\to\infty}\int_{\{\tau>n\}}|X_n|d\mathbf{P}=0,$$

但是却没有 $\mathbf{E}\,|X_\tau|<\infty$ (意即 $\mathbf{E}\,|X_\tau|=\infty$).

17. 设 $M=(M_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 是鞅, 而 $\tau_N=\inf\{n\geqslant 0:\ |M_n|\geqslant N\},\quad N\geqslant 1$ $(\inf\varnothing=\infty)$. 证明, 鞅 M 一致可积, 当且仅当

$$\lim_N\mathbf{E}\,|M_{\tau_N}|I(\tau_N<\infty)=0.$$

18. (续例 1 和例 2) 设 $\xi_1,\xi_2,\cdots$ 是相互独立的对称伯努利随机变量序列 (即有 $\mathbf{P}\{\xi_i=1\}=\mathbf{P}\{\xi_i=-1\}=1/2,\ i\geqslant 1$). 我们来考察停时 $\tau=\inf\{n\geqslant 0:\ S_n=1\}$, 其中 $S_0=0,\ S_n=\xi_1+\cdots+\xi_n,\ n\geqslant 1$, 且 $\inf\varnothing=\infty$.

(a) 对 $\lambda\in\mathbb{R}$, 令

$$X_n^\lambda=\frac{e^{\lambda S_n}}{(\mathrm{ch}\lambda)^n}.$$

证明, 对任何 $\lambda\in\mathbb{R}$, 序列 $(X_n^\lambda)_{n\geqslant 0}$ 都是鞅. 并利用这一事实, 证明, $\mathbf{P}\{\tau<\infty\}=1$, $\mathbf{E}\tau=\infty$, 且对于一切 $\lambda\geqslant 0$, 都有

$$\mathbf{E}(\mathrm{ch}\lambda)^{\tau}=e^{-\lambda a}$$

(试比较 (18) 式).

(b) 令 $\alpha=1/\mathrm{ch}\lambda$, 由所得的公式知

$$\mathbf{E}\,\alpha^\tau=\sum_{n\geqslant 1}\alpha^n\mathbf{P}\{\tau=n\}=\frac{1}{\alpha}(1-\sqrt{1-\alpha^2}).$$

(亦可参阅第八章第 8 节第 26 题.) 由此得知

$$\mathbf{P}\{\tau=2n-1\}=(-1)^{n+1}C_{1/2}^n,$$

其中

$$C_X^n=\frac{X(X-1)\cdots(X-n+1)}{n!}$$

(参阅第一章第 2 节第 22 题).

(c) 令 $I = \inf\{S_n:\ n \leqslant \tau\}$. 证明, 对一切 $k \geqslant 0$, 都有

$$\mathbf{P}\{I \leqslant -k\} = \frac{1}{k+1}.$$

(d) 令 $\tau_k = \inf\{n \geqslant 0:\ S_n = 1 \text{ 或 } S_n = -k\}$. 证明, $\tau_k \to \tau$ ($\mathbf{P}$-a.s.), $S_{\tau_k} \to S_\tau$ ($\mathbf{P}$-a.s.), $k \to \infty$, 但是 $\mathbf{E}\,S_{\tau_k} \not\to \mathbf{E}\,S_\tau$ ($\mathbf{E}\,S_{\tau_k} = 0$, $\mathbf{E}\,S_\tau = 1$). 试说明 "此处尽管有 $S_{\tau_k} \to S_\tau$ ($\mathbf{P}$-a.s.), $k \to \infty$", 但是 $\mathbf{E}\,S_{\tau_k}$ 却不收敛到 $\mathbf{E}\,S_\tau$ 的原因.

19. 在定理 2 和 3 的条件中包含有所考察的马尔可夫时 τ 具有有限期望的假设 (即 $\mathbf{E}\,\tau < \infty$). 试证明, 若能找到整数 N 和 $0 < \varepsilon < 1$, 使得对一切 $n \geqslant 1$, 都有

$$\mathbf{P}\{\tau \leqslant n + N \,|\, \mathscr{F}_n\} > \varepsilon \quad (\mathbf{P}\text{-a.s.}),$$

则有 $\mathbf{E}\,\tau < \infty$.

提示: 用归纳法证明,

$$\mathbf{P}\{\tau \geqslant kN\} \leqslant (1-\varepsilon)^k, \quad k \geqslant 1.$$

20. 设 $m(t)$ 为更新函数 (参阅第 4 小节). 初等更新理论断言: $m(t)/t \to 1/\mu$, $t \to \infty$. 如下两个命题可以视为对这一断言的精确化.

(a) 如果更新过程 $N = (N_t)_{t\geqslant 0}$ 是 d-格子点的 (对于 $d > 0$, 随机变量 σ_1 的分布函数的支撑是集合 $\{0, d, 2d, \cdots\}$), 则有 (柯尔莫戈洛夫, 1936)

$$\sum_{k=1}^{\infty} \mathbf{P}\{T_k = nd\} \to \frac{d}{\mu}, \quad n \to \infty.$$

(b) 如果更新过程 $N = (N_t)_{t\geqslant 0}$ 对任何 $d > 0$, 都不是 d-格子点的, 那么对任何 $h > 0$, 都有 (布莱克韦尔, 1948)

$$\sum_{k=1}^{\infty} \mathbf{P}\{t < T_k \leqslant t + h\} \to \frac{h}{\mu}, \quad n \to \infty. \qquad (*)$$

($(*)$ 式的左端等于 $m(t+h) - m(t)$.)

提示: 为证 (a), 最好仔细研究第八章第 6 节定理 2 的证明.

21. 设 $\xi_1, \xi_2, \cdots$ 为相互独立的随机变量序列, 有 $\mathbf{P}\{\xi_i = 1\} = p$, $\mathbf{P}\{\xi_i = -1\} = q$, $p+q = 1$, $i \geqslant 1$. 设 x, a, b 都是整数, 其中 $a \leqslant 0 \leqslant b$. 令 $S_n(x) = x + \xi_1 + \cdots + \xi_n$ 和

$$\tau_a(x) = \inf\{n \geqslant 0:\ S_n(x) \leqslant a\},$$
$$\tau^b(x) = \inf\{n \geqslant 0:\ S_n(x) \geqslant b\},$$
$$\tau_a^b(x) = \inf\{n \geqslant 0:\ S_n(x) \leqslant a \text{ 或 } S_n(x) \geqslant b\}.$$

证明,

$$\mathbf{P}\{\tau_a(x) < \infty\} = \begin{cases} 1, & \text{如果 } p \leqslant q \text{ 且 } x > a, \\ \left(\dfrac{q}{p}\right)^{x-a}, & \text{如果 } p > q \text{ 且 } x > a, \end{cases}$$

$$\mathbf{P}\{\tau^b(x) < \infty\} = \begin{cases} 1, & \text{如果 } p \geqslant q \text{ 且 } x < b, \\ \left(\dfrac{p}{q}\right)^{b-x}, & \text{如果 } p < q \text{ 且 } x < b, \end{cases}$$

$$\mathbf{P}\{\tau_a^b(x) < \infty\} = 1, \quad a \leqslant x \leqslant b.$$

并且对于 $a \leqslant x \leqslant b$, 有

$$\mathbf{E}\tau_a^b(x) = \frac{x-a}{q-p} - \frac{b-a}{q-p}\left(\frac{(q/p)^x - (q/p)^a}{(q/p)^b - (q/p)^a}\right), \quad \text{如果 } p \neq q,$$

$$\mathbf{E}\tau_a^b(x) = (b-a)(x-a), \quad \text{如果 } p = q = 1/2.$$

22. 设 $\xi_1, \xi_2, \cdots$ 为独立同分布的随机变量序列, 它们的取值集合为 $\{-1, 0, 1, \cdots\}$, 且具有中位数 $\mu < 0$. 令 $S_0 = 1$, $S_n = 1 + \xi_1 + \cdots + \xi_n$, $n \geqslant 1$ 及 $\tau = \inf\{n \geqslant 0:\ S_n = 0\}$. 证明, $\mathbf{E}\tau = \frac{1}{|\mu|}$.

§3. 一些基本不等式

1. 设 $X = (X_n, \mathscr{F}_n)_{n\geqslant 0}$ 是非负半鞅, $V = (V_n, \mathscr{F}_{n-1})_{n\geqslant 0}$ 是可预报序列 ($\mathscr{F}_{-1} = \mathscr{F}_0$), 有 $0 \leqslant V_{n+1} \leqslant V_n \leqslant C$ ($\mathbf{P}$-a.s.), 其中 C 为常数. 证明, 可将不等式 (1) 推广为如下形式: 对任何 $\lambda > 0$, 都有

$$\lambda \mathbf{P}\left\{\max_{0\leqslant k\leqslant n} V_k X_k \geqslant \lambda\right\} + \int\limits_{\left\{\max\limits_{0\leqslant k\leqslant n} V_k X_k < \lambda\right\}} V_n X_n \, d\mathbf{P} \leqslant \sum_{k=0}^{n} \mathbf{E}\, V_k \Delta X_k \quad (\Delta X_0 = X_0).$$

2. 证明下述克里克伯格 (Krickbeg) 分解的正确性: 任何满足条件 $\sup \mathbf{E}\,|X_n| < \infty$ 的鞅 $X = (X_n, \mathscr{F}_n)_{n\geqslant 0}$ 都可以表示为两个非负鞅的差.

3. 设 $\xi_1, \xi_2, \cdots$ 为相互独立的随机变量序列. 令 $S_n = \xi_1 + \cdots + \xi_n$ 和 $S_{m,n} = \sum\limits_{j=m+1}^{n} \xi_j$. 证明如下的奥坦韦安尼 (Ottaviani) 不等式:

$$\mathbf{P}\left\{\max_{1\leqslant j\leqslant n} |S_j| > 2t\right\} \leqslant \frac{\mathbf{P}\{|S_n| > t\}}{\min\limits_{1\leqslant j\leqslant n} \mathbf{P}\{|S_{j,n}| \leqslant t\}}, \quad t > 0.$$

再在条件 $\mathbf{E}\,\xi_i = 0\ (i \geqslant 1)$ 之下, 由上式推出:

$$\int_0^\infty \mathbf{P}\left\{\max_{1\leqslant j\leqslant n} |S_j| > 2t\right\} dt \leqslant 2\mathbf{E}\,|S_n| + 2\int_{2\mathbf{E}\,|S_n|}^{\infty} \mathbf{P}\{|S_n| > t\}\, dt. \qquad (*)$$

提示: 为证奥坦韦安尼不等式, 应当令 $A = \left\{\max_{1\leqslant k\leqslant n} |S_k| \geqslant 2t\right\}$,

$$A_k = \{|S_i| < 2t,\ i = 1, \cdots, k-1;\ |S_k| \geqslant 2t\}, \quad 1 \leqslant k \leqslant n,$$

则有 $A = \sum_{k=1}^{n} A_k$, 并可断言, 对 $t > 0$, 有

$$\begin{aligned}\mathbf{P}(A)\left[\min_{1\leqslant j\leqslant n} \mathbf{P}\{|S_{j,n}| \leqslant t\}\right] &= (\mathbf{P}(A_1) + \cdots + \mathbf{P}(A_n))\left[\min_{1\leqslant j\leqslant n} \mathbf{P}\{|S_{j,n}| \leqslant t\}\right] \\ &\leqslant \mathbf{P}(A_1 \cap \{|S_n| > t\}) + \cdots + \mathbf{P}(A_n \cap \{|S_n| > t\}) \leqslant \mathbf{P}\{|S_n| > t\}.\end{aligned}$$

为证不等式 (*), 只需指出, 在条件 $\mathbf{E}\,\xi_i = 0\,(i \geqslant 1)$ 之下, 有

$$\int_0^\infty \mathbf{P}\left\{\max_{1\leqslant j\leqslant n} |S_j| > 2t\right\} dt \leqslant \int_0^{2\mathbf{E}\,|S_n|} dt + \int_{2\mathbf{E}\,|S_n|}^\infty \frac{\mathbf{P}\{|S_n| > t\}}{1 - \max_{1\leqslant j\leqslant n} \mathbf{P}\{|S_{j,n}| > t\}}\, dt$$

以及在 $t \geqslant 2\mathbf{E}\,|S_n|$ 时, 有

$$\begin{aligned}1 - \max_{1\leqslant j\leqslant n} \mathbf{P}\{|S_{j,n}| > t\} &\geqslant 1 - \max_{1\leqslant j\leqslant n} \mathbf{P}\{|S_{j,n}| > 2\mathbf{E}\,|S_n|\} \\ &\geqslant 1 - \max_{1\leqslant j\leqslant n} \frac{\mathbf{E}\,|S_{j,n}|}{2\mathbf{E}\,|S_n|} \geqslant 1 - \frac{1}{2} = \frac{1}{2}.\end{aligned}$$

4. 设 $\xi_1, \xi_2, \cdots$ 为相互独立的随机变量序列, 有 $\mathbf{E}\,\xi_i = 0$. 试利用第 3 题中的不等式 (*), 证明, 可以将不等式 (10) 加强为

$$\mathbf{E}\max_{1\leqslant j\leqslant n} |S_j| \leqslant 8\mathbf{E}\,|S_n|.$$

提示: 应当利用上一题中的不等式 (*).

5. 试证明 (16) 式.

6. 试证明不等式 (19).

7. 设 $\mathscr{F}_0, \cdots, \mathscr{F}_n$ 为 σ-代数, 有 $\mathscr{F}_0 \subseteq \mathscr{F}_1 \subseteq \cdots \subseteq \mathscr{F}_n$. 设事件 $A_k \in \mathscr{F}_k$, $k = 1, \cdots, n$. 试利用 (22) 式证明如下的德沃里茨基 (Dvoretzky) 不等式: 对 $\lambda > 0$, 有

$$\mathbf{P}\left\{\bigcup_{k=1}^{n} A_k\right\} \leqslant \lambda + \mathbf{P}\left\{\sum_{k=1}^{n} \mathbf{P}(A_k \,|\, \mathscr{F}_{k-1}) > \lambda\right\}.$$

提示: 应当定义 $X_k = I(A_k)$, $k = 1, \cdots, n$, 因而有 $X_n^* = \max_{1\leqslant k\leqslant n} I(A_k) = I\left(\bigcup_{k=1}^{n} A_k\right)$. 如果记 $Y_n = \sum_{k=1}^{n} \mathbf{P}(A_k \,|\, \mathscr{F}_{k-1})$, 则由 (22) 式, 得

$$\mathbf{P}\{X_N^* \geqslant 1\} \leqslant \mathbf{E}(Y_n \wedge \varepsilon) + \mathbf{P}\{Y_n \geqslant \varepsilon\},$$

由此即可得出所要证明的不等式.

8. 设 $X = (X_n)_{n\geqslant 1}$ 是平方可积鞅, $(b_n)_{n\geqslant 1}$ 是非降的正数序列. 证明如下的哈伊克–雷尼 (Hajek-Renyi) 不等式:

$$\mathbf{P}\left\{\max_{1\leqslant k\leqslant n} \left|\frac{X_k}{b_k}\right| \geqslant \lambda\right\} \leqslant \frac{1}{\lambda^2} \sum_{k=1}^{n} \frac{\mathbf{E}\,(\Delta X_k^2)}{b_n^2}, \quad \Delta X_k = X_k - X_{k-1}, \quad X_0 = 0.$$

9. 设 $X = (X_n)_{n\geqslant 1}$ 是半鞅, $g(x)$ 是上升的非负下凸函数. 证明, 对任何正数 h 和任何实数 x, 都有

$$\mathbf{P}\left\{\max_{1\leqslant k\leqslant n} X_k \geqslant x\right\} \leqslant \frac{\mathbf{E}\,g(hX_n)}{g(hx)}.$$

特别地, 有如下的杜布指数型不等式:

$$\mathbf{P}\left\{\max_{1\leqslant k\leqslant n} X_k \geqslant x\right\} \leqslant e^{-hx}\mathbf{E}\,e^{hX_n}.$$

(试比较第四章第 2 节第 23 题中的柯尔莫戈洛夫指数型不等式.)

10. 设 $\xi_1, \xi_2, \cdots$ 是相互独立的随机变量序列, 有 $\mathbf{E}\,\xi_n = 0$, $\mathbf{E}\,\xi_n^2 = 1$, $n \geqslant 1$. 令 $\tau = \inf\left\{n \geqslant 1 : \sum_{i=1}^{n} \xi_i > 0\right\}$, 其中 $\inf \varnothing = \infty$. 证明, $\mathbf{E}\,\tau^{1/2} < \infty$.

11. 设 $\xi = (\xi_n)_{n\geqslant 1}$ 是鞅差序列, $1 < p \leqslant 2$. 证明,

$$\mathbf{E}\sup_{n\geqslant 1}\left|\sum_{i=1}^{n} \xi_i\right|^p \leqslant C_p \sum_{i=1}^{\infty} \mathbf{E}\,|X_i|^p,$$

其中 C_p 是某个常数.

12. 设 $X = (X_k)_{k\geqslant 1}$ 是鞅, 有 $\mathbf{E}\,X_k = 0$, $\mathbf{E}\,X_k^2 < \infty$. (作为第四章第 2 节第 5 题中的不等式的推广) 证明, 对一切 $n \geqslant 1$, 都有

$$\mathbf{P}\left\{\max_{1\leqslant k\leqslant n} X_k \geqslant \varepsilon\right\} \leqslant \frac{\mathbf{E}\,X_n^2}{\varepsilon^2 + \mathbf{E}\,X_n^2}, \quad \varepsilon \geqslant 0.$$

13. 设 $\xi_1, \xi_2, \cdots$ 是相互独立的随机变量序列, 有 $\mathbf{P}\{\xi_n = 1\} = p$, $\mathbf{P}\{\xi_n = -1\} = q$, $p + q = 1$, $0 < p < 1$. 令 $S_0 = 0$, $S_n = \xi_1 + \cdots + \xi_n$, $n \geqslant 1$.

(a) 证明, 则序列 $((q/p)^{S_n})_{n\geqslant 0}$ 是鞅; 又如果 $p < q$, 则有如下的最大值不等式:

$$\mathbf{P}\left\{\sup_{n\geqslant 0} S_n \geqslant k\right\} \leqslant \left(\frac{p}{q}\right)^k$$

(如果 $p \geqslant q$, 该不等式显然成立).

(b) 证明, 如果 $p < q$, 则有

$$\mathbf{E}\sup_{n\geqslant 0} S_n \leqslant \frac{p}{q-p}.$$

(c) 证明, 上述两个不等式中的等号都是成立的, 即当 $p < q$ 时, 随机变量 $\sup_{n\geqslant 0} S_n$ 具有几何分布 (参阅第二章第 3 节表 2):

$$\mathbf{P}\left\{\sup_{n\geqslant 0} S_n = k\right\} = \left(\frac{p}{q}\right)^k \left(1-\frac{p}{q}\right), \quad k=0,1,2,\cdots,$$

因而 $\mathbf{E}\sup_{n\geqslant 0} S_n = \dfrac{p}{q-p}$.

14. 设 $M=(M_k,\ \mathscr{F}_k)_{0\leqslant k\leqslant n}$ 是鞅, $M_0=0$, 并且 $-a_k \leqslant \Delta M_k \leqslant 1-a_k$, $k=1,\cdots,n$, 其中 $\Delta M_k = M_k - M_{k-1}$, $a_k\in[0,1]$. 作为第四章第 5 节第 14 题中结论的推广, 证明, 对于一切 $0\leqslant x<q$, 其中 $q=1-p$, $p=\frac{1}{n}\sum_{k=1}^n a_k$, 都有如下的不等式成立:

$$\mathbf{P}\{M_n \geqslant nx\} \leqslant e^{n\psi(x)},$$

其中 $\psi(x) = \ln\left[\left(\frac{p}{p+x}\right)^{p+x}\left(\frac{q}{q-x}\right)^{q-x}\right]$.

提示: 参阅第四章第 5 节第 14 题提示, 证明, 在现在的情形下, 有

$$\begin{aligned}\mathbf{E}\, e^{hM_n} &= \mathbf{E}\left[e^{hM_{n-1}}\mathbf{E}\left(e^{h\Delta M_n} \mid \mathscr{F}_{n-1}\right)\right] \\ &\leqslant \mathbf{E}\left[e^{hM_{n-1}}\left((1-a_n)e^{-ha_n}+a_n e^{h(1-a_n)}\right)\right].\end{aligned}$$

15. 设 $M=(M_k,\ \mathscr{F}_k)_{k\geqslant 0}$ 是鞅, 有 $M_0=0$, 并且对非负常数 a_k 和 b_k, 有

$$a_k \leqslant \Delta M_k \leqslant b_k, \quad k\geqslant 1,$$

其中 $\Delta M_k = M_k - M_{k-1}$.

(a) 证明, 对一切 $x\geqslant 0$ 和所有 $n\geqslant 1$, 都有

$$\mathbf{P}\{M_n \geqslant x\} \leqslant \exp\left\{-\frac{2x^2}{\sum_{k=1}^n (a_k+b_k)^2}\right\},$$

$$\mathbf{P}\{M_n \leqslant -x\} \leqslant \exp\left\{-\frac{2x^2}{\sum_{k=1}^n (a_k+b_k)^2}\right\},$$

因此亦有

$$\mathbf{P}\{|M_n| \geqslant x\} \leqslant 2\exp\left\{-\frac{2x^2}{\sum_{k=1}^n (a_k+b_k)^2}\right\}.$$

(试比较第一章第 6 节和第四章第 5 节中的相应不等式.)

(b) 证明, 如果对一切 $k\geqslant 1$, 都有 $a_k=a$, $b_k=b$ (此即意味着对一切 $k\geqslant 1$, 都有 $a\leqslant \Delta M_k \leqslant b$), 则有如下的最大值不等式: 对 $\beta>0$ 和 $x>0$, 都有

$$\mathbf{P}\{M_n - \beta n \geqslant x \text{ 对某个 } n\} \leqslant \exp\left\{-\frac{8x\beta}{(a+b)^2}\right\}; \tag{$*$}$$

而对于 $\beta>0$ 和整数 $m\geqslant 1$, 都有

$$\mathbf{P}\{M_n \geqslant \beta n \text{ 对某个 } n\geqslant m\} \leqslant \exp\left\{-\frac{2m\beta^2}{(a+b)^2}\right\},$$

$$\mathbf{P}\{M_n \leqslant -\beta n \text{ 对某个 } n\geqslant m\} \leqslant \exp\left\{-\frac{2m\beta^2}{(a+b)^2}\right\}.$$

(试比较第四章第 5 节中的不等式.)

注: 通常将 (a) 中的不等式称为霍夫丁-阿佐姆 (Hoeffding-Azym) 不等式, 它的推广形式 (b) 在参考文献 [101] 中给出.

提示: (a) 对于任何 $c>0$, 都有

$$\mathbf{P}\{M_n \geqslant x\} \leqslant e^{-cx}\mathbf{E}\, e^{cM_n}.$$

记 $V_n = e^{cM_n}$, 则有 $V_n = V_{n-1}e^{c\Delta M_n}$, 这就表明

$$\mathbf{E}\,(V_n|M_{n-1}) = V_{n-1}\mathbf{E}\,(e^{c\Delta M_n}|M_{n-1}).$$

通过对 n 迭代, 并利用条件 $-a_k\leqslant \Delta M_k \leqslant b_k$, 得到

$$\mathbf{P}\{M_n \geqslant x\} \leqslant e^{-cx}\prod_{k=1}^n \frac{b_k e^{-ca_k}+a_k e^{-cb_k}}{a_k+b_k} \leqslant e^{-cx}\prod_{k=1}^n \exp\left\{\frac{c^2}{8}(a_k+b_k)^2\right\},$$

亦即

$$\mathbf{P}\{M_n \geqslant x\} \leqslant \exp\left\{-cx + c^2\sum_{k=1}^n \frac{(a_k+b_k)^2}{8}\right\}.$$

再由 $c>0$ 的任意性, 即得

$$\mathbf{P}\{M_n \geqslant x\} \leqslant \min_{c>0}\exp\left\{-cx+c^2\sum_{k=1}^n\frac{(a_k+b_k)^2}{8}\right\} = \exp\left\{-\frac{2x^2}{\sum_{k=1}^n (a_k+b_k)^2}\right\}.$$

(b) 为证 $(*)$, 需要引入变量

$$V_n = \exp\{c(M_n - x - \beta n)\}, \quad n\geqslant 0.$$

应当指出, 若取 $c=\frac{8\beta}{(a+b)^2}$, 则序列 $(V_n)_{n\geqslant 1}$ 形成非负上鞅, 它对于任何有限的马尔可夫时 $\tau(K)$ $(\leqslant K)$, 都有

$$\mathbf{E}\, V_{\tau(K)} \leqslant \mathbf{E}\, V_0 = e^{-8x\beta/(a+b)^2}.$$

令 $\tau(K)=\min\{n: M_n \geqslant x+\beta n$ 或 $n=K\}$, 得到

$$\mathbf{P}\{M_{\tau(K)} \geqslant x+\beta\tau(K)\}=\mathbf{P}\{V_{\tau(K)}\geqslant 1\}\leqslant \mathbf{E}\,V_{\tau(K)}\leqslant \mathbf{E}\,V_0,$$

亦即

$$\mathbf{P}\{M_n \geqslant x+\beta n \text{ 对某个 } n\leqslant K\}\leqslant \exp\{-\frac{8x\beta}{(a+b)^2}\}.$$

令 $K\to\infty$, 即得不等式 $(*)$. 由它按如下方式推导, 可得进一步的不等式 $(**)$:

$$\begin{aligned}\mathbf{P}\{M_n\geqslant \beta n \text{ 对某个 } n\geqslant m\} &\leqslant \mathbf{P}\left\{M_n\geqslant \frac{m\beta}{2}+\frac{n\beta}{2} \text{ 对某个 } n\right\}\\ &\leqslant \exp\left\{-\frac{8(m\beta/2)(\beta/2)}{(a+b)^2}\right\}=\exp\left\{-\frac{2m\beta^2}{(a+b)^2}\right\}.\end{aligned}$$

16. 设 $M=(M_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 是鞅, 而 $\tau=\inf\{n\geqslant 0:\ |M_n|\geqslant\lambda\}$, 其中 $\lambda>0$ 和 $\inf\varnothing=\infty$. 证明,

$$\mathbf{P}\{\tau<\infty\}\leqslant \lambda^{-1}\|M\|_1,$$

$\|M\|_1=\sup\limits_n \mathbf{E}|M_n|$.

17. 符号同上题, 证明,

$$\sum_{k=0}^{\infty}\mathbf{E}|M_k-M_{k-1}|^2 I(\tau>k)\leqslant 2\|M\|_1,$$

其中, $M_{-1}=0$.

18. 设 $M=(M_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 是鞅, 其中 $M_0=0$. 令 $\Delta M_k=M_k-M_{k-1}$, $[M]_n=\sum\limits_{k=1}^{n}(\Delta M_k)^2$, 记 $[M]=([M]_n,\ \mathscr{F}_n)_{n\geqslant 1}$, 并将其称为 $M=(M_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 的二次变型. 证明, 如下二条件

$$\mathbf{E}\sup_n |M_n|<\infty \quad 与 \quad \mathbf{E}\,[M]_\infty^{1/2}<\infty$$

是相互等价的:

$$\mathbf{E}\sup_n |M_n|<\infty \quad\Leftrightarrow\quad \mathbf{E}\,[M]_\infty^{1/2}<\infty. \qquad (*)$$

(与此有关的还应提及伯克霍尔德–戴维斯–甘地不等式:

$$A_p\|[M]_\infty^{1/2}\|_p\leqslant \|M_\infty^*\|_p\leqslant B_p\|[M]_\infty^{1/2}\|_p,\quad p\geqslant 1,$$

其中, $M_\infty^*=\sup\limits_n |M_n|$, A_p 与 B_p 均为仅与 p 有关的常数 (参阅 (27), (30), 以及专著 [72]), 可将该不等式视为上面的不等式 $(*)$ 的 "L^p 精确化".)

19. 设 $M=(M_k,\ \mathscr{F}_k)_{k\geqslant 1}$ 是鞅. 证明如下的伯克霍尔德不等式: 对任何 $r\geqslant 2$, 存在仅依赖于 r 的常数 B_r, 使得

$$\mathbf{E}\,|M_n|^r\leqslant B_r\left\{\mathbf{E}\left(\sum_{k=1}^{n}\mathbf{E}\left((\Delta M_k)^2\,|\,\mathscr{F}_{k-1}\right)\right)^{r/2}+\mathbf{E}\sup_{1\leqslant k\leqslant n}|\Delta M_k|^r\right\},$$

其中 $\Delta M_k=M_k-M_{k-1}$, $k\geqslant 1$, 而 $M_0=0$.

20. (**矩不等式 I**) 设 $\xi_1,\xi_2,\cdots$ 是独立同分布的随机变量序列, 有 $\mathbf{E}\,\xi_1=0$, 且对某个 $r\geqslant 1$, 有 $\mathbf{E}\,|\xi_1|^r<\infty$. 令 $S_n=\xi_1+\cdots+\xi_n$, $n\geqslant 1$. 则根据马钦凯维奇–济格蒙德不等式 (26) 的右半部, 可以得到:

$$\mathbf{E}\,|S_n|^r\leqslant B_r\mathbf{E}\left(\sum_{i=1}^{n}\xi_i^2\right)^{r/2},$$

其中 B_r 是某个仅依赖于 r 的常数.

试运用关于 $r\geqslant 2$ 的闵可夫斯基 (Minkowski) 不等式 (第二章第 6 节) 和关于 $r<2$ 的 c_r 不等式 (第二章第 6 节第 72 题), 证明,

$$\mathbf{E}\,|S_n|^r\leqslant\begin{cases}B_r n\mathbf{E}\,|\xi_1|^r, & 1\leqslant r\leqslant 2,\\ B_r n^{r/2}\mathbf{E}\,|\xi_1|^r, & r>2.\end{cases}$$

(特别地, 对于 $r\geqslant 2$, 可由该式得到

$$\mathbf{E}\,|n^{-1/2}S_n|^r\leqslant B_r\mathbf{E}\,|\xi_1|^r.$$

根据[①] 第三章第 4 节第 22 题, 有 $\lim\limits_n \mathbf{E}\,|n^{-1/2}S_n|^r=\mathbf{E}\,|Z|^r$, 其中 $Z\sim\mathcal{N}(0,\sigma^2)$, $\sigma^2=\mathbf{E}\,\xi_1^2$.)

21. (**矩不等式 II**) 设 $\xi_1,\xi_2,\cdots$ 是独立同分布的随机变量序列, $S_0=0$, $S_n=\xi_1+\cdots+\xi_n$, $n\geqslant 1$. 而 τ 是一个关于滤子 $(\mathscr{F}_n^S)_{n\geqslant 0}$ 的停时, 其中 $\mathscr{F}_0^S=\{\varnothing,\Omega\}$, $\mathscr{F}_n^S=\sigma\{S_1,\cdots,S_n\}$. 证明,

(a) 如果 $0<r\leqslant 1$, 且 $\mathbf{E}\,|\xi_1|^r<\infty$, 则

$$\mathbf{E}\,|S_\tau|^r\leqslant \mathbf{E}\,|\xi_1|^r\mathbf{E}\,\tau.$$

(b) 如果 $1<r\leqslant 2$, 且 $\mathbf{E}\,|\xi_1|^r<\infty$, $\mathbf{E}\,\xi_1=0$, 则

$$\mathbf{E}\,|S_\tau|^r\leqslant B_r\mathbf{E}\,|\xi_1|^r\mathbf{E}\,\tau.$$

①原书此处有多处笔误: 将 "第三章" 错为 "第二章", 将 "$\mathbf{E}\,|n^{-1/2}S_n|^r$" 错为 "$\mathbf{E}\,n^{-1/2}|S_n|^r$", 将 "$\lim\limits_n \mathbf{E}\,|n^{-1/2}S_n|^r=\mathbf{E}\,|Z|^r$" 错为 "$\lim\limits_n \mathbf{E}\,n^{-1/2}|S_n|^r\to\mathbf{E}\,|Z|^r$", 等等 —— 译者注.

(c) 如果 $r>2$, 且 $\mathbf{E}|\xi_1|^r<\infty$, $\mathbf{E}\,\xi_1=0$, 则

$$\mathbf{E}\,|S_\tau|^r\leqslant B_r[(\mathbf{E}\,\xi_1^2)^{r/2}\mathbf{E}\,\tau^{r/2}+\mathbf{E}\,|\xi_1|^r\mathbf{E}\,\tau]\leqslant 2B_r\mathbf{E}\,|\xi_1|^r\mathbf{E}\,\tau^{r/2},$$

其中 B_r 是仅与 r 有关的常数.

提示: 在每种情况下, 都应首先考虑 "截尾的"(有限的) 停时 $\tau\wedge n$ $(n\geqslant 1)$, 然后再令 $n\to\infty$.

22. 设 $\xi_1,\cdots,\xi_n$ 是相互独立的随机变量, 试证明如下形式的马钦凯维奇–济格蒙德不等式: 对 $r>0$ 和任何 $n\geqslant 1$, 都存在常数 $B=B(r)$(试比较不等式 (26)), 使得

$$\mathbf{E}\left|\sum_{j=1}^n\xi_j\right|^{2r}\leqslant Bn^{r-1}\sum_{j=1}^n\mathbf{E}\,|\xi_j|^{2r}.$$

提示: 只需讨论整数 $r\geqslant 1$ 的情形, 该种情况下的证明较为简单.

23. 设 $(\xi_n)_{n\geqslant 1}$ 是正交随机变量序列 ($\mathbf{E}\,\xi_i\xi_j=0,\ i\neq j;\ \mathbf{E}\,\xi_i^2=1,$ 一切 $i\geqslant 1$). 证明, 对于任意实数序列 $(c_n)_{n\geqslant 1}$, 都有如下的拉德马赫–梅尼绍夫 (Rademacher-Men'shov) 不等式成立: 对任何 $n\geqslant 1$

$$\mathbf{E}\max_{1\leqslant k\leqslant n}\left(\sum_{j=1}^k c_j\xi_j\right)^2\leqslant \ln^2(4n)\sum_{j=1}^n c_j^2.$$

24. 设 $(\xi_n)_{n\geqslant 1}$ 是正交随机变量序列, $(c_n)_{n\geqslant 1}$ 是实数序列, 有

$$\sum_{k=1}^\infty c_k^2\ln^2 k<\infty.$$

则级数 $\sum\limits_{k=1}^\infty c_k\xi_k$ 概率为 1 地收敛.

提示: 利用上题中的结论.

25. (**关于伯努利随机变量的极值作用. I.**) 设 $\xi_1,\cdots,\xi_n$ 是相互独立的伯努利随机变量, 有 $\mathbf{P}\{\xi_i=1\}=\mathbf{P}\{\xi_i=-1\}=1/2$.

(a) 证明, 由关于 $p=2m$, $m\geqslant 1$ 情形的辛钦不等式 (26)(参阅《概率》第二卷 p137) 的右半部可以推出: 对任何 $n\geqslant 1$, 都有

$$\mathbf{E}\left|\sum_{k=1}^n a_k\xi_k\right|^{2m}\leqslant\mathbf{E}\left|\sum_{k=1}^n a_kX_k\right|^{2m},$$

其中 $X_1,\cdots,X_n$ 是相互独立的标准正态 $\mathcal{N}(0,1)$ 随机变量.

(b) 以 Σ_n 表示具有性质 $\mathbf{D}\,X_i=1,\ i=1,\cdots,n$, 的独立同分布的对称随机变量 $X_1,\cdots,X_n$ 的类. 证明, 对任何 $n\geqslant 1$ 和 $m\geqslant 1$, 都有

$$\mathbf{E}\left|\sum_{k=1}^n a_k\xi_k\right|^{2m}\leqslant\inf_{(X_1,\cdots,X_n)\in\Sigma_n}\mathbf{E}\left|\sum_{k=1}^n a_kX_k\right|^{2m}.$$

提示: (a) 只需证明

$$\mathbf{E}\left|\sum_{k=1}^n a_k\xi_k\right|^{2m}=\sum_{k_1+\cdots+k_n=m,\ k_i\geqslant 0}\frac{(2m)!}{(2k_1)!\cdots(2k_n)!}|a_1|^{2k_1}\cdots|a_n|^{2k_n},$$

$$\mathbf{E}\left|\sum_{k=1}^n a_kX_k\right|^{2m}=\frac{(2m)!}{2^m m!}\sum_{k_1+\cdots+k_n=m,\ k_i\geqslant 0}\frac{m!}{k_1!\cdots k_n!}|a_1|^{2k_1}\cdots|a_n|^{2k_n},$$

以及当 $k_1+\cdots+k_n=m,\ k_i\geqslant 0$ 时, 有 $2^mk_1!\cdots k_n!\leqslant(2k_1)!\cdots(2k_n)!$. (注意, $\frac{(2m)!}{2^m m!}=(2m-1)!!=\mathbf{E}\,X_1^{2m}$, 参阅第二章第 8 节第 9 题.)

(b) 当 $m=1$ 时, 所要证明的不等式显然成立. 设 $m\geqslant 2$, 此时 $\varphi(t)=\mathbf{E}\,|x+\sqrt{t}\xi_1|^{2m}$ 是 $t>0$ 的凸函数, 再通过对条件期望使用延森不等式, 得到

$$\mathbf{E}\left|\sum_{k=1}^n a_k\xi_k\right|^{2m}\leqslant\mathbf{E}\left|\sum_{k=1}^n a_k\xi_k|X_k|\right|^{2m},$$

其中假定 $\xi_1,\cdots,\xi_n$ 与 $X_1,\cdots,X_n$ 相互独立. 最后只需注意

$$(\xi_1|X_1|,\cdots,\xi_n|X_n|)\overset{\text{law}}{=}(X_1,\cdots,X_n).$$

26. (**关于伯努利随机变量的极值作用. II.**) 设 $X_1,\cdots,X_n$ 是相互独立的随机变量, 有 $\mathbf{P}\{0\leqslant X_i\leqslant 1\}=1,\ \mathbf{E}\,X_i=p_i,\ i=1,\cdots,n$. 设 $\xi_1,\cdots,\xi_n$ 是独立同分布的伯努利随机变量, 有 $\mathbf{P}\{\xi_i=1\}=p,\ \mathbf{P}\{\xi_i=0\}=1-p$, 其中 $p=\frac{p_1+\cdots+p_n}{n}$. 证明如下的本特古斯 (Bentcus) 不等式: 对一切 $x=0,1,2,\cdots$

$$\mathbf{P}\{X_1+\cdots+X_n\geqslant x\}\leqslant e\mathbf{P}\{\xi_1+\cdots+\xi_n\geqslant x\},$$

其中 $e=2.718\cdots$, $n\geqslant 1$.

§4. 半鞅和鞅收敛的基本定理

1. 设 $(\mathscr{G}_n)_{n\geqslant 1}$ 是非升的 σ-代数序列, $\mathscr{G}_1\supseteq\mathscr{G}_2\supseteq\cdots$, $\mathscr{G}_\infty=\bigcap\mathscr{G}_n$, 而 η 是某个可积的随机变量. 证明如下的与定理 3 相类似的结论: 当 $n\to\infty$ 时, 有

$$\mathbf{E}\,(\eta\,|\,\mathscr{G}_n)\to\mathbf{E}\,(\eta\,|\,\mathscr{G}_\infty)\quad(\mathbf{P}\text{-a.s. 和 }L^1).$$

提示: 令 $M_k=\mathbf{E}\,(\eta\,|\,\mathscr{G}_k)$, 以 $\beta_n(a,b)$ 表示序列 $M=(M_k)_{1\leqslant k\leqslant n}$ "从上往下穿越区间 (a,b) 的次数". 再令 $\beta_\infty(a,b)=\lim\limits_n\beta_n(a,b)$, 并证明

$$\mathbf{E}\,\beta_\infty(a,b)\leqslant\frac{\mathbf{E}\,|\eta|+|a|}{b-a}<\infty,$$

于是得知 $\beta_\infty(a,b)<\infty$ ($\mathbf{P}$-a.s.). 接下来的步骤与定理 1 和定理 3 的证明类似.

2. 设 $\xi_1, \xi_2, \cdots$ 是独立同分布随机变量序列, 有 $\mathbf{E}|\xi_1| < \infty$ 和 $\mathbf{E}\xi_1 = m$. 令 $S_n = \xi_1 + \cdots + \xi_n$, $n \geqslant 1$. 证明 (参阅第二章第 7 节第 2 题),

$$\mathbf{E}(\xi_1 \mid S_n, S_{n+1}, \cdots) = \mathbf{E}(\xi_1 \mid S_n) = \frac{S_n}{n} \quad (\mathbf{P}\text{-a.s.}),$$

并由前一题推出强大数律: 当 $n \to \infty$ 时, 有

$$\frac{S_n}{n} \to m \quad (\mathbf{P}\text{-a.s. 和 } L^1).$$

提示: 可以证明对每个 $B \in \sigma(S_n, S_{n+1}, \cdots)$, 都有 $\mathbf{E}I_B\xi_i = \mathbf{E}I_B\xi_1$, $i \leqslant n$, 由此即得

$$\mathbf{E}(S_n \mid S_n, S_{n+1}, \cdots) = n\mathbf{E}(\xi_1 \mid S_n, S_{n+1}, \cdots),$$

从而 $\mathbf{E}(\xi_1 \mid S_n, S_{n+1}, \cdots) = \frac{S_n}{n}$ ($\mathbf{P}$-a.s.). 为证 "$\frac{S_n}{n} \to m$ ($\mathbf{P}$-a.s. 和 L^1)", 需要引入 σ-代数 $\mathscr{X}(S) = \bigcap_{n=1}^{\infty} \sigma(S_n, S_{n+1}, \cdots)$, 由第 1 题的结论可以得到 $\frac{S_n}{n} \to \mathbf{E}(\xi_1 \mid \mathscr{X}(S))$ ($\mathbf{P}$-a.s. 和 L^1). 最后, 再对 $\mathscr{X}(S)$ 中的事件 A 利用休伊特与塞维杰的 "0-1 律"(参阅第四章第 1 节定理 3).

3. 如下的结果是共同使用勒贝格控制收敛定理与莱维定理的结果, 试证明其正确性: 设 $(\xi_n)_{n\geqslant 1}$ 是随机变量序列, 有 $\xi_n \to \xi$ ($\mathbf{P}$-a.s.), $|\xi_n| \leqslant \eta$, $\mathbf{E}\eta < \infty$, 而 $\{\mathscr{F}_m, m \geqslant 1\}$ 是非降的 σ-族, $\mathscr{F}_\infty = \sigma(\bigcup \mathscr{F}_m)$. 则 $\mathbf{P}$-a.s. 地有

$$\lim_{m\to\infty,\ n\to\infty} \mathbf{E}(\xi_n \mid \mathscr{F}_m) = \mathbf{E}(\xi \mid \mathscr{F}_\infty).$$

提示: 写

$$\begin{aligned}&\mathbf{E}(\xi_n \mid \mathscr{F}_m) - \mathbf{E}(\xi \mid \mathscr{F}_\infty)\\ &\quad = [\mathbf{E}(\xi_n \mid \mathscr{F}_m) - \mathbf{E}(\xi \mid \mathscr{F}_m)] + [\mathbf{E}(\xi \mid \mathscr{F}_m) - \mathbf{E}(\xi \mid \mathscr{F}_\infty)],\end{aligned}$$

对充分大的 n 和 m, 利用勒贝格控制收敛定理 (第二章第 6 节定理 3) 估计上式右端第一个括弧; 利用莱维定理 (定理 3) 估计上式右端第二个括弧.

4. 证明 (12) 式.

提示: 对于所有关于 $\mathscr{F}_n = \sigma(H_1, \cdots, H_n)$ 可测的函数而言, 函数族 $\{H_1(x), \cdots, H_n(x)\}$ 是一组基. 由于 $\mathscr{F}_n$ 中的成员数目有限, 所以关于它可测的函数都是简单函数 (第二章第 4 节引理 3), 这表明带有有限个常数 $a_1, \cdots, a_n$ 的 (12) 式成立. 于是由基 $\{H_1(x), \cdots, H_n(x)\}$ 的正交性, 知 $a_k = (f, H_k)$.

5. 设 $\Omega = [0,1)$, $\mathscr{F} = \mathscr{B}([0,1))$, $\mathbf{P}$ 为勒贝格测度, 而 $f = f(x) \in L^1$. 令

$$f_n(x) = 2^n \int_{k2^{-n}}^{(k+1)2^{-n}} f(y)dy, \quad \text{如果 } k2^{-n} \leqslant x < (k+1)2^{-n}.$$

证明, $f_n(x) \to f(x)$ ($\mathbf{P}$-a.s.).

提示: 令 $\mathscr{F}_n = \sigma([j2^{-n}, (j+1)2^{-n}),\ j = 0, 1, \cdots, 2^n - 1)$. 证明中的关键步骤是证明 $(f_n(x), \mathscr{F}_n)_{n\geqslant 1}$ 是鞅. 接下来再利用定理 1.

6. 设 $\Omega = [0,1)$, $\mathscr{F} = \mathscr{B}([0,1))$, $\mathbf{P}$ 为勒贝格测度, 而 $f = f(x) \in L^1$. 假定该函数在区间 $[0,2)$ 上为周期函数, 令

$$f_n(x) = \sum_{i=1}^{2^n} 2^{-n} f(x + i2^{-n}).$$

证明,

$$f_n(x) \to f(x) \quad (\mathbf{P}\text{-a.s.}).$$

提示: 定义与上题类似的 σ-代数, 并证明 $(f_n(x), \mathscr{F}_n)_{n\geqslant 1}$ 是鞅, 再利用定理 1.

7. 设 $X = (X_n, \mathscr{F}_n)_{n\geqslant 1}$ 为广义上鞅, 即有

$$\inf_m \sup_{n\geqslant m} \mathbf{E}(X_n^+ \mid \mathscr{F}_m) < \infty \quad (\mathbf{P}\text{-a.s.}).$$

证明, 定理 1 对于广义上鞅仍然成立.

8. 设 $(a_n)_{n\geqslant 1}$ 为某个数列, 对所有满足条件 $|t| < \delta$, $\delta > 0$ 的实数 t, 都有极限 $\lim_{n\to\infty} e^{ita_n}$ 存在. 证明, 极限 $\lim_{n\to\infty} a_n$ 存在有限.

提示: "极限 $\lim_{n\to\infty} e^{ita_n}$, $|t| < \delta$, 存在" 等价于 "对一切 $t \in \mathbb{R}$, 都有极限 $\lim_{n\to\infty} e^{ita_n}$ 存在". 故只需证明, 对某个常数 c, 可将极限函数 $f(t) = \lim_{n\to\infty} e^{ita_n}$ 表示为 e^{itc} 的形式. 为此, 可以通过证明 $f(t)$ 具有如下各条性质来实现:

(i) $|f(t)| = 1$, $t \in \mathbb{R}$.

(ii) $f(t_1 + t_2) = f(t_1) + f(t_2)$, $t_1, t_2 \in \mathbb{R}$.

(iii) $f(t)$ 的连续点集在 $\mathbb{R}$ 中处处稠密.

9. 设 $F = F(x)$, $x \in \mathbb{R}$, 是分布函数, $\alpha \in (0,1)$. 假设存在 $\theta \in \mathbb{R}$, 使得 $F(\theta) = \alpha$. 我们来构造序列 (罗宾斯–蒙罗 (Robbins-Monro) 程序) $X_1, X_2, \cdots$ 如下:

$$X_{n+1} = X_n - n^{-1}(Y_n - \alpha), \quad X_0 = 0,$$

其中 $Y_1, Y_2, \cdots$ 是随机变量, 有

$$\mathbf{P}\{Y_1 = y\} = \begin{cases} F(0), & \text{如果 } y = 1, \\ 1 - F(0), & \text{如果 } y = 0, \end{cases}$$

而对于 $n > 1$, 则有

$$\mathbf{P}\{Y_n = y \mid X_1, \cdots, X_n;\ Y_1, \cdots, Y_{n-1}\} = \begin{cases} F(X_n), & \text{如果 } y = 1, \\ 1 - F(X_n), & \text{如果 } y = 0. \end{cases}$$

证明 "随机逼近论" 中的如下结果: 对于罗宾斯–蒙罗程序, 有 $\mathbf{E}|X_n-\theta|^2 \to 0,\ n\to\infty$.

10. 设 $X=(X_n,\mathscr{F}_n)_{n\geqslant 1}$ 是半鞅, 对任何停时 τ, 都有 $\mathbf{E}(X_\tau I(\tau<\infty))\neq\infty$. 证明, 概率为 1 地存在极限 $\lim\limits_{n\to\infty} X_n$.

11. 设 $X=(X_n,\mathscr{F}_n)_{n\geqslant 1}$ 是鞅, $\mathscr{F}_\infty=\sigma\left(\bigcup\limits_{n=1}^{\infty}\mathscr{F}_n\right)$. 证明, 如果序列 $(X_n)_{n\geqslant 1}$ 一致可积, 则极限 $X_\infty=\lim\limits_{n\to\infty}X_n$ ($\mathbf{P}$-a.s.) 存在, 并且 "闭" 序列 $\overline{X}=(X_n,\mathscr{F}_n)_{1\leqslant n\leqslant\infty}$ 是鞅.

12. 现在设 $X=(X_n,\mathscr{F}_n)_{n\geqslant 1}$ 是半鞅, $\mathscr{F}_\infty=\sigma\left(\bigcup\limits_{n=1}^{\infty}\mathscr{F}_n\right)$. 证明, 如果序列 $(X_n^+)_{n\geqslant 1}$ 一致可积, 则极限 $X_\infty=\lim\limits_{n\to\infty}X_n$ ($\mathbf{P}$-a.s.) 存在, 并且 "闭" 序列 $\overline{X}=(X_n,\mathscr{F}_n)_{1\leqslant n\leqslant\infty}$ 是半鞅.

13. 由定理 1 的推论 1, 可以推出: 如果 X 是非负上鞅, 则概率为 1 地存在有限极限 $X_\infty=\lim\limits_{n\to\infty}X_n$. 证明, 由此还可以进一步推出如下各条性质:

(a) $\mathbf{E}(X_\infty\,|\,\mathscr{F}_n)\leqslant X_n$ ($\mathbf{P}$-a.s.), $n\geqslant 1$.

(b) $\mathbf{E}X_\infty\leqslant\varliminf\limits_{n\to\infty}\mathbf{E}X_n$.

(c) 对任何停时 τ 与 σ, 都有 $\mathbf{E}(X_\tau\,|\,\mathscr{F}_\sigma)\leqslant X_{\tau\wedge\sigma}$.

(d) 对一切满足条件 $\frac{g(x)}{x}\to 0,\ x\to\infty$ 的连续函数 $g=g(x),\ x\geqslant 0$, 都有 $\mathbf{E}g(X_\infty)=\lim\limits_{n\to\infty}\mathbf{E}g(X_n)$.

(e) 如果对一切 $x>0$, 都有 $g(x)>g(0)=0$, 则有

$$X_\infty=0 \quad\Leftrightarrow\quad \lim_{n\to\infty}\mathbf{E}g(X_n)=0.$$

(f) 对任何给定的 $0<p<1$, 都有

$$\mathbf{P}\{X_\infty=0\}=1 \quad\Leftrightarrow\quad \lim_{n\to\infty}\mathbf{E}X_n^p=0.$$

14. 在莱维定理 (定理 3) 中假定了条件 $\mathbf{E}|\xi|<\infty$ 成立. 试举例说明, 为了使得 $\mathbf{E}(\xi\,|\,\mathscr{F}_n)\to\mathbf{E}(\xi\,|\,\mathscr{F})$ ($\mathbf{P}$-a.s.), 可以仅要求 $\mathbf{E}\xi$ 确定 (意即有 $\min(\mathbf{E}\xi^+,\mathbf{E}\xi^-)<\infty$), 而不需假定 $\mathbf{E}|\xi|<\infty$.

15. 如果 $X=(X_n,\mathscr{F}_n)_{n\geqslant 1}$ 是鞅, 满足条件 $\sup\limits_n\mathbf{E}|X_n|<\infty$, 则概率为 1 地存在 (定理 1) 极限 $\lim\limits_{n\to\infty}X_n$. 试举例说明, 存在鞅 X, 虽然对其有 $\sup\limits_n\mathbf{E}|X_n|=\infty$, 但却依然概率为 1 地存在着极限 $\lim\limits_{n\to\infty}X_n$.

16. 试举例说明, 存在鞅 $(X_n)_{n\geqslant 0}$, 对其有 $X_n\to-\infty,\ n\to\infty$.

17. 对于一致可积的半鞅 (上鞅) $X=(X_n,\mathscr{F}_n)_{n\geqslant 1}$, (根据第 4 节定理 2) 存在 "术语上的" 随机变量 X_∞, 使得 $X_n\to X_\infty$ ($\mathbf{P}$-a.s.). 试举例说明, 存在半鞅 (上鞅) 序列, 对其有 "术语上的" 随机变量 X_∞ 存在, 但是序列 $(X_n)_{n\geqslant 1}$ 却不是一致可积的.

18. 证明, 满足条件

$$\sup_n\mathbf{E}\left(|X_n|\ln^+|X_n|\right)<\infty$$

的鞅 $X=(X_n)_{n\geqslant 0}$ 是莱维鞅.

19. 试举例说明, 存在非负鞅 $X=(X_n,\mathscr{F}_n)_{n\geqslant 1}$, 对每个 $n\geqslant 1$, 都有 $\mathbf{E}X_n=1$, 而且当 $n\to\infty$ 时, 逐点地有 $X_n(\omega)\to 0$, 但是却有 $\mathbf{E}\sup\limits_n X_n=\infty$.

20. 设 $X=(X_n,\mathscr{F}_n)_{n\geqslant 1}$ 是一致可积的半鞅. 证明, 对于一切马尔可夫时 τ, 都有

$$\mathbf{E}(X_\infty\,|\,\mathscr{F}_\tau)\geqslant X_\tau \quad (\mathbf{P}\text{-a.s.}).$$

(随机变量 X_∞ 就是, 根据第 12 题, 以概率 1 存在的极限 $\lim\limits_n X_n$.)

21. (续定理 1) 试举例说明, 存在半鞅 $X=(X_n,\mathscr{F}_n)_{n\geqslant 1}$, 对其有 $\sup\limits_n\mathbf{E}|X_n|<\infty$, 因而概率为 1 地存在 $\lim\limits_n X_n$ $(=X_\infty)$, 但是却有 $X_n\overset{L^1}{\nrightarrow}X_\infty$.

22. 设 $M=(M_n,\mathscr{F}_n)_{n\geqslant 1}$ 是平方可积鞅. 证明, 条件

$$\sum_{k\geqslant 1}\mathbf{E}(M_k-M_{k-1})^2<\infty$$

(亦即条件 $\mathbf{E}\langle M\rangle_\infty<\infty$, 其中 $\langle M\rangle_\infty=\lim\limits_n\langle M\rangle_n$) 既可保证 $M_n\to M_\infty$ ($\mathbf{P}$-a.s.), 又可保证 $M_n\overset{L^1}{\to}M_\infty$, 其中 M_∞ 是某个满足条件 $\mathbf{E}M_\infty^2<\infty$ 的随机变量.

23. 在定义半鞅 $X=(X_n,\mathscr{F}_n)_{n\geqslant 0}$ 时, 假定了对一切 $n\geqslant 0$, 都有 $\mathbf{E}|X_n|<\infty$. 其实这一条件经常被削弱为 $\mathbf{E}X_n^-<\infty,\ n\geqslant 0$. 事实上, 半鞅的许多性质在这个关于 $X_n,\ n\geqslant 0$ 的可积性的削弱条件之下, 仍然得以保持.

试逐一检验第 2 ~ 4 小节中所列举的半鞅的性质中, 哪些性质在削弱后的可积性条件之下仍可成立.

24. 若 $X=(X_n,\mathscr{F}_n)_{n\geqslant 0}$ 是上鞅 (即 X_n 为 $\mathscr{F}_n$ 可测, $\mathbf{E}|X_n|<\infty$, 且 $\mathbf{E}(X_{n+1}\,|\mathscr{F}_n)\leqslant X_n,\ n\geqslant 0$), 则根据定理 1, 在条件 $\sup\limits_n\mathbf{E}|X_n|<\infty$ 之下, 概率为 1 地存在极限 $\lim\limits_n X_n=X_\infty$ 且 $\mathbf{E}|X_\infty|<\infty$.

我们指出, 半鞅条件 "$\mathbf{E}(X_{n+1}\,|\,\mathscr{F}_n)\leqslant X_n$" 即使在没有 "$\mathbf{E}|X_{n+1}|<\infty$" 的假定条件之下也依然可以成立. 数学期望 $\mathbf{E}(X_{n+1}\,|\,\mathscr{F}_n)$, 例如在条件 $X_{n+1}\geqslant 0$ 之下, 就可以被确定, 虽然其值可能为无穷.

与此相关联的, 如果对一切 $n\geqslant 0$, 都有 $\mathbf{P}\{X_n\geqslant 0\}=1$ 及 $\mathbf{E}(X_{n+1}\,|\,\mathscr{F}_n)\leqslant X_n$ ($\mathbf{P}$-a.s.), 则称此时的 $X=(X_n,\mathscr{F}_n)_{n\geqslant 0}$ 是非负上鞅序列.

证明, 对于非负上鞅序列 $X=(X_n,\mathscr{F}_n)_{n\geqslant 0}$, 概率为 1 地存在极限 $\lim\limits_n X_n$ $(=X_\infty)$. 并且如果 $\mathbf{P}\{X_0<\infty\}=1$, 则有 $\mathbf{P}\{X_\infty<\infty\}=1$.

提示: 利用建立在关于相交数的第 3 节定理 5 中的 (37) 式基础上的定理 1 的证明.

25. (续第二章第 2 节第 15 题) 第二章第 2 节第 15 题曾经指出: 关于 σ-代数的如下等式一般来说是不成立的:

$$\bigcap_n \sigma(\mathscr{G},\ \mathscr{E}_n)=\sigma(\mathscr{G},\ \bigcap_n \mathscr{E}_n).$$

试证明, 为了该等式能够成立, 只需满足条件:

对于一切 $n>1$, σ-代数 $\mathscr{G}$ 与 $\mathscr{E}_1$ 都关于 σ-代数 $\mathscr{E}_n$ 相对独立, 亦即对 $A\in\mathscr{G}$, $B\in\mathscr{E}_1$, 都 ($\mathbf{P}$-a.s. 地) 有

$$\mathbf{P}\{A\cap B\,|\,\mathscr{E}_n\}=\mathbf{P}\{A\,|\,\mathscr{E}_n\}\mathbf{P}\{B\,|\,\mathscr{E}_n\}.$$

提示: 只需证明, 对每个 $\mathscr{G}\vee\mathscr{E}_1$ 可测的有界随机变量 X, 都 ($\mathbf{P}$-a.s. 地) 有

$$\mathbf{E}\left(X\ |\ \bigcap_n(\mathscr{G}\vee\mathscr{E}_n)\right)=\mathbf{E}\left(X\ |\ (\mathscr{G}\vee\bigcap_n\mathscr{E}_n)\right).$$

而为此, 又只需对如下形式的随机变量证明这一结论:

$$X=X_1X_2,$$

其中 X_1 与 X_2 皆为有界随机变量, X_1 为 $\mathscr{E}_1$ 可测, X_2 为 $\mathscr{E}_2$ 可测.

然后再利用已经证得的条件独立性及第 1 题中的 L^1 收敛性.

26. 设 $\xi_1,\xi_2,\cdots$ 是相互独立的非负随机变量序列, 有 $\mathbf{E}\xi_i\leqslant 1$ 和 $\mathbf{P}\{\xi_i=1\}<1$①. 令 $M_n=\xi_1\cdots\xi_n$, $n\geqslant 1$. 证明, 当 $n\to\infty$ 时, 有 $M_n\to 0$ ($\mathbf{P}$-a.s.).

提示: 注意序列 $(M_n)_{n\geqslant 1}$ 是非负上鞅.

27. 设 $(\Omega,(\mathscr{F}_i)_{i\geqslant 0},\mathbf{P})$ 是可滤空间, 其中 $\mathscr{F}_0=\{\Omega,\ \varnothing\}$. 设 $\xi_1,\xi_2,\cdots$ 是随机变量序列, 其中 ξ_i 是 $\mathscr{F}_i$ 可测的. 假设有 $\sup\limits_i\mathbf{E}\,|\xi_i|^\alpha<\infty$, 其中 $\alpha\in(1,2]$. 证明,

$$\frac{1}{n}\sum_{i=1}^n(\xi_i-\mathbf{E}\,(\xi_i\,|\,\mathscr{F}_{i-1}))\overset{\text{a.s.与 }L^\alpha}{\longrightarrow}0.$$

(试比较第四章第 3 节中的强大数定律和第五章第 3 节中的遍历性定理.)

§5. 半鞅和鞅的收敛集

1. 证明, 如果半鞅 $X=(X_n,\ \mathscr{F}_n)_{n\geqslant 1}$ 满足条件 $\mathbf{E}\sup\limits_n|X_n|<\infty$, 则它属于 $\mathbf{C}^+$ 类.

①原书中为 "有 $\mathbf{E}\xi_1\leqslant 1$ 和 $\mathbf{P}\{\xi_1=1\}<1$" —— 译者注.

2. 证明, 定理 1 和定理 2 对于广义半鞅仍然成立.

3. 证明, 对于广义半鞅, ($\mathbf{P}$-a.s. 地) 成立如下结论:

$$\left\{\inf_m\sup_{n\geqslant m}\mathbf{E}\,(X_n^+\,|\,\mathscr{F}_m)<\infty\right\}\subseteq\{X_n\ \to\}.$$

4. 证明, 定理 1 的推论对于广义半鞅仍然成立.

5. 证明, $\mathbf{C}^+$ 类中的任何广义半鞅都是局部半鞅.

6. 设 $a_n>0$, $n\geqslant 1$, 令 $b_n=\sum\limits_{k=1}^n a_k$. 证明, $\sum\limits_{n=1}^\infty\dfrac{a_n}{b_n^2}<\infty$.

提示: 分 $\sum\limits_{n=1}^\infty a_n<\infty$ 与 $\sum\limits_{n=1}^\infty a_n=\infty$ 两种情况讨论.

7. 设 $\xi_0,\xi_1,\xi_2,\cdots$ 为一致有界的随机变量序列: $|\xi_n|\leqslant c$, $n\geqslant 0$. 证明, 级数 $\sum\limits_{n\geqslant 0}\xi_n$ 与级数 $\sum\limits_{n\geqslant 1}\mathbf{E}\,(\xi_n\,|\,\xi_0,\xi_1,\cdots,\xi_{n-1})$ 具有相同的 ($\mathbf{P}$-a.s.) 敛散性, 即同为 ($\mathbf{P}$-a.s.) 收敛或同为 ($\mathbf{P}$-a.s.) 发散.

8. 设 $X=(X_n)_{n\geqslant 0}$ 是鞅, 对某个 $c<\infty$, 有 $\Delta X_n=X_n-X_{n-1}\leqslant c$ ($\mathbf{P}$-a.s.) ($\Delta X_0=X_0$). 证明, ($\mathbf{P}$-a.s. 地) 有

$$\{X_n\ \to\}=\left\{\sup_n X_n<\infty\right\}.$$

9. 设 $X=(X_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 是鞅, 满足条件 $\sup\limits_n\mathbf{E}|X_n|<\infty$. 证明, $\sum\limits_{n\geqslant 1}(\Delta X_n)^2<\infty$ ($\mathbf{P}$-a.s.). (试比较第 3 节第 18 题.)

10. 设 $X=(X_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 是鞅, 满足条件 $\mathbf{E}\sup\limits_{n\geqslant 1}|\Delta X_n|<\infty$. 证明,

$$\left\{\sum_{n\geqslant 1}(\Delta X_n)^2<\infty\right\}\subseteq\{X_n\ \to\}.$$

特别地, 如果 $\mathbf{E}\left(\sum\limits_{n\geqslant 1}(\Delta X_n)^2\right)^{1/2}<\infty$, 则序列 $X=(X_n)_{n\geqslant 0}$ 概率为 1 地收敛.

11. 设 $X=(X_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 是鞅, 满足条件 $\sup\limits_n\mathbf{E}|X_n|<\infty$; 而设 $Y=(Y_n)_{n\geqslant 1}$ 是可预报序列, 满足条件 $\sup\limits_{n\geqslant 1}|Y_n|<\infty$ ($\mathbf{P}$-a.s.). 证明, 级数 $\sum\limits_{n\geqslant 1}Y_n\Delta X_n$ ($\mathbf{P}$-a.s.) 收敛.

12. 设 $X=(X_n,\ \mathscr{F}_n)_{n\geqslant 0}$ 是鞅, 有 $\sup\mathbf{E}\,(|\Delta X_\tau|I(\tau<\infty))<\infty$, 其中的上确界系对所有有限的正停时 τ 所取. 证明,

$$\left\{\sum_{n\geqslant 1}(\Delta X_n)^2<\infty\right\}\subseteq\{X_n\ \to\infty\}.$$

13. 设 $M=(M_n, \mathscr{F}_n)$ 是平方可积鞅. 证明, 对集合 $\{\langle M\rangle_\infty=\infty\}$ 中的几乎所有 ω, 都有

$$\lim_n \frac{M_n}{\langle M\rangle_n}=0.$$

§6. 概率测度在带滤子的可测空间上的绝对连续性和奇异性

1. 试证明 (6) 式.

2. 设 $\widetilde{\mathbf{P}}_n \sim \mathbf{P}_n$, $n\geqslant 1$. 证明

$$\widetilde{\mathbf{P}} \sim \mathbf{P} \quad \Leftrightarrow \quad \widetilde{\mathbf{P}}\{z_\infty<\infty\}=\mathbf{P}\{z_\infty>0\}=1,$$
$$\widetilde{\mathbf{P}} \perp \mathbf{P} \quad \Leftrightarrow \quad \widetilde{\mathbf{P}}\{z_\infty=\infty\}=1 \text{ 或 } \mathbf{P}\{z_\infty=0\}=1.$$

3. 设 $\widetilde{\mathbf{P}}_n \ll \mathbf{P}_n$, $n\geqslant 1$, 而 τ 是 (关于 $(\mathscr{F}_n)$ 的) 停时, $\widetilde{\mathbf{P}}_\tau=\widetilde{\mathbf{P}}|\mathscr{F}_\tau$ 与 $\mathbf{P}_\tau=\mathbf{P}|\mathscr{F}_\tau$ 分别是测度 $\widetilde{\mathbf{P}}$ 与 $\mathbf{P}$ 在 σ-代数 $\mathscr{F}_\tau$ 上的限制. 证明, 如果 $\{\tau=\infty\}=\{z_\infty<\infty\}$ ($\mathbf{P}$-a.s.), 则有 $\widetilde{\mathbf{P}}_\tau \ll \mathbf{P}_\tau$. (特别地, 如果 $\mathbf{P}\{\tau<\infty\}=1$, 则有 $\widetilde{\mathbf{P}}_\tau \ll \mathbf{P}_\tau$.)

4. 试分别证明计数公式 (21) 和 (22).

提示: 直接验证对任何 $A\in\mathscr{F}_{n-1}$, 都有

$$\mathbf{E}\,[I_A\widetilde{\mathbf{E}}\,(\eta|\mathscr{F}_{n-1})z_{n-1}]=\mathbf{E}\,[I_A\eta z_n].$$

为证明第二个公式, 应当注意: $\widetilde{\mathbf{P}}\{z_{n-1}=0\}=0$.

5. 试证明不等式 (28), (29) 和 (32).

6. 试证明 (34) 式.

7. 设第 2 小节中的序列 $\xi=(\xi_1,\xi_2,\cdots)$ 与 $\widetilde{\xi}=(\widetilde{\xi}_1,\widetilde{\xi}_2,\cdots)$ 均由独立同分布的随机变量构成.

(a) 证明, 如果 $P_{\widetilde{\xi}_1}\ll P_{\xi_1}$, 则 $\widetilde{P}\ll P$, 当且仅当, 测度 $P_{\widetilde{\xi}_1}$ 与 P_{ξ_1} 相互重合. 如果 $P_{\widetilde{\xi}_1}\ll P_{\xi_1}$ 且 $P_{\widetilde{\xi}_1}\neq P_{\xi_1}$, 则 $\widetilde{P}\perp P$.

(b) 证明, 若 $P_{\widetilde{\xi}_1}\sim P_{\xi_1}$, 则二者不可兼得: 或者 $\widetilde{P}=P$, 或者 $\widetilde{P}\perp P$ (试比较定理 3 中的角谷 "不可兼得说").

8. 设 $\mathbf{P}$ 与 $\widetilde{\mathbf{P}}$ 都是可滤空间 $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geqslant 1})$ 上的概率测度. 设 $\widetilde{\mathbf{P}} \overset{\mathrm{loc}}{\ll} \mathbf{P}$ (即对所有 $n\geqslant 1$, 都有 $\widetilde{\mathbf{P}}_n\ll\mathbf{P}_n$, 其中 $\mathbf{P}_n=\mathbf{P}|\mathscr{F}_n$, $\widetilde{\mathbf{P}}_n=\widetilde{\mathbf{P}}|\mathscr{F}_n$.) 令 $z_n=\frac{d\widetilde{\mathbf{P}}_n}{d\mathbf{P}_n}$, $n\geqslant 1$.

证明, 如果 τ 是马尔可夫时, 则在集合 $\{\tau<\infty\}$ 上 ($\mathbf{P}$-a.s.地) 有

$$\widetilde{\mathbf{P}}_\tau\ll\mathbf{P}_\tau \quad \text{与} \quad \frac{d\widetilde{\mathbf{P}}_\tau}{d\mathbf{P}_\tau}=z_\tau.$$

9. 证明, $\widetilde{\mathbf{P}} \overset{\mathrm{loc}}{\ll} \mathbf{P}$, 当且仅当, 存在具有性质 $\mathbf{P}\{\lim\limits_n \tau_n=\infty\}=1$ 的上升的停时序列 $(\tau_n)_{n\geqslant 1}$, 使得对每个 $n\geqslant 1$, 都有 $\widetilde{\mathbf{P}}_{\tau_n}\ll\mathbf{P}_{\tau_n}$.

10. 设 $\widetilde{\mathbf{P}} \overset{\mathrm{loc}}{\ll} \mathbf{P}$, 令 $z_n=\frac{d\widetilde{\mathbf{P}}_n}{d\mathbf{P}_n}$, $n\geqslant 1$. 证明,

$$\widetilde{\mathbf{P}}\left\{\inf_n z_n>0\right\}=1.$$

11. 设 $\widetilde{\mathbf{P}} \overset{\mathrm{loc}}{\ll} \mathbf{P}$, 令 $z_n=\frac{d\widetilde{\mathbf{P}}_n}{d\mathbf{P}_n}$, $n\geqslant 1$, $\mathscr{F}_\infty=\sigma(\bigcup\mathscr{F}_n)$.

证明下列各条件相互等价:

(i) $\widetilde{\mathbf{P}}_\infty\ll\mathbf{P}_\infty$, 其中, $\widetilde{\mathbf{P}}_\infty=\widetilde{\mathbf{P}}|\mathscr{F}_\infty$, 而 $\mathbf{P}_\infty=\mathbf{P}|\mathscr{F}_\infty$.

(ii) $\widetilde{\mathbf{P}}\left\{\sup\limits_n z_n<\infty\right\}=1$.

(iii) 鞅 $(z_n, \mathscr{F}_n)_{n\geqslant 1}$ 一致可积.

12. 设 $(\Omega, \mathscr{F}, \mathbf{P})$ 为概率空间, $\mathscr{G}$ 是 $\mathscr{F}$ 的可分的子 σ-代数, 它是由 $\mathscr{F}$ 的子集类 $\{G_n:\ n\geqslant 1\}$ 所生成的. 令 $\mathscr{G}_n=\sigma(G_1,\cdots,G_n)$, 而 $\mathscr{D}_n$ 是生成 $\mathscr{G}_n$ 的关于 Ω 的最小分割.

设 $\mathbf{Q}$ 是 $(\Omega, \mathscr{F})$ 上的另一个概率测度. 令

$$X_n(\omega)=\sum_{A\in\mathscr{D}_n}\frac{\mathbf{Q}(A)}{\mathbf{P}(A)}I_A(\omega)$$

(约定 $0/0=0$). 证明,

(a) 序列 $(X_n, \mathscr{G}_n)_{n\geqslant 1}$ 是 (关于 $\mathbf{P}$ 的) 上鞅.

(b) 如果 $\mathbf{Q}\ll\mathbf{P}$, 则序列 $(X_n, \mathscr{G}_n)_{n\geqslant 1}$ 是鞅.

13. (续上题) 证明, 如果 $\mathbf{Q}\ll\mathbf{P}$, 则存在 $\mathscr{G}$ 可测的随机变量 $X_\infty=X_\infty(\omega)$, 使得 $X_n\xrightarrow{L^1}X_\infty$, $X_n=\mathbf{E}\,(X_\infty\,|\,\mathscr{G}_n)$ ($\mathbf{P}$-a.s.), 并且对于任何 $A\in\mathscr{G}$, 都有

$$\mathbf{Q}(A)=\int_A X_\infty d\mathbf{P}.$$

(本结论不是别的, 它恰恰就是关于可分的子 σ-代数 $\mathscr{G}\subseteq\mathscr{F}$ 的拉东–尼科迪姆定理 (第二章第 6 节).)

14. (续角谷 "不可兼得说") 设 $\alpha_1,\alpha_2,\cdots$ 是相互独立的非负随机变量序列, 有 $\mathbf{E}\,\alpha_i=1$. 令 $z_n=\prod\limits_{k=1}^n\alpha_k$, $z_0=1$. 证明,

(a) 序列 $(z_n)_{n\geqslant 0}$ 是 (非负) 鞅.

(b) 概率为 1 地存在极限 $\lim\limits_n z_n$ $(=z_\infty)$.

(c) 下列各条件相互等价:

(i) $\mathbf{E}\,z_\infty=1$. (ii) $z_n\xrightarrow{L^1}z_\infty$. (iii) $(z_n)_{n\geqslant 0}$ 为一致可积族.

(iv) $\sum\limits_{n=1}^\infty(1-\mathbf{E}\sqrt{\alpha_n})<\infty$. (v) $\prod\limits_{n=1}^\infty\mathbf{E}\sqrt{\alpha_n}>0$.

§7. 随机游动越出曲线边界的概率的渐近式

1. 证明, 由 (4) 式所定义的序列是鞅. 如果取消条件 $|\alpha_n| \leqslant c$ ($\mathbf{P}$-a.s.), $n \geqslant 1$, 那么结论是否还成立?

2. 试证明 (13) 式.

提示: 只需将 $\mathbf{E}\, z_n^p$ 表示为:

$$\mathbf{E}\, z_n^p = \prod_{k=2}^{n}\left(p\exp\left\{\alpha_k\xi_k - \frac{1}{2}\alpha_k^2\right\}\right),$$

再利用随机变量 ξ_k 的正态性 ($\xi_k \sim \mathcal{N}(0,1)$).

3. 试证明 (17) 式.

4. 设 $\xi_1, \xi_2, \cdots$ 是独立同分布的随机变量序列, $S_n = \xi_1 + \cdots + \xi_n$, $n \geqslant 1$. 对 $c > 0$, 令

$$\tau_{(\geqslant 0)} = \inf\{n \geqslant 1:\ S_n \geqslant 0\}, \quad \tau_{(>c)} = \inf\{n \geqslant 1:\ S_n > c\}$$

(如同往常, $\inf \varnothing = \infty$). 证明,

(a) $\mathbf{P}\{\tau_{(\geqslant 0)} < \infty\} = 1 \Leftrightarrow \mathbf{P}\{\overline{\lim} S_n = \infty\} = 1$.

(b) $(\mathbf{E}\,\tau_{(\geqslant 0)} < \infty) \Leftrightarrow$ (对一切 $c > 0$, $\mathbf{E}\,\tau_{(>c)} < \infty$).

5. 在上题的基础上, 再定义

$$\tau_{(>0)} = \inf\{n \geqslant 1:\ S_n > 0\}, \quad \tau_{(\leqslant 0)} = \inf\{n \geqslant 1:\ S_n \leqslant 0\},$$

$$\tau_{(<0)} = \inf\{n \geqslant 1:\ S_n < 0\}.$$

证明,

$$\mathbf{E}\,\tau_{(\geqslant 0)} = \frac{1}{\mathbf{P}\{\tau_{(<0)} = \infty\}} \quad \text{与} \quad \mathbf{E}\,\tau_{(>0)} = \frac{1}{\mathbf{P}\{\tau_{(\leqslant 0)} = \infty\}}.$$

6. 设 $\xi_1, \xi_2, \cdots$ 是独立同分布的随机变量序列, 有 $\mathbf{E}|\xi_1| > 0$. 令 $S_n = \xi_1 + \cdots + \xi_n$, $n \geqslant 1$, 并设 $\frac{S_n}{n} \xrightarrow{\mathbf{P}} 0$.

证明, 如上所定义的马尔可夫时 $\tau_{(\geqslant 0)}$ 与 $\tau_{(\leqslant 0)}$ 都是有限的, 并且概率为 1 地有 $\overline{\lim} S_n = \infty$, $\underline{\lim} S_n = -\infty$.

7. 在上题的条件下, 证明, 概率为 1 地有 $S_n \to \infty$, 当且仅当, 存在 (关于流 $\mathscr{F}^\xi = (\mathscr{F}_n^\xi)_{n\geqslant 1}$, 其中 $\mathscr{F}_n^\xi = \sigma(\xi_1, \cdots, \xi_n)$ 的) 停时 τ, 使得 $\mathbf{E}\,\tau < \infty$ 和 $\mathbf{E}\, S_\tau > 0$.

8. 设 $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geqslant 0}, \mathbf{P})$ 是可滤的概率空间. 令

$$h_n = \mu_n + \sigma_n \xi_n, \quad n \geqslant 1,$$

其中, μ_n 与 σ_n 是 $\mathscr{F}_{n-1}$ 可测的, $\sigma_n > 0$. 而 $\xi = (\xi_n, \mathscr{F}_n)_{n\geqslant 1}$ 是相互独立的标准正态 $\mathcal{N}(0,1)$ 随机变量序列. 证明, 序列 $h = (h_n, \mathscr{F}_n)_{n\geqslant 1}$ 是条件高斯的, 即

$$\mathrm{Law}(h_n \,|\, \mathscr{F}_{n-1}; \mathbf{P}) = \mathcal{N}(\mu_n, \sigma_n^2) \quad (\mathbf{P}\text{-a.s.}).$$

令

$$Z_n = \exp\left\{-\sum_{k=1}^{n}\frac{\mu_k}{\sigma_k}\xi_k - \frac{1}{2}\sum_{k=1}^{n}\left(\frac{\mu_k}{\sigma_k}\right)^2\right\}, \quad n \geqslant 1.$$

证明,

(a) 序列 $Z_n = (Z_n)_{n\geqslant 1}$ (关于测度 $\mathbf{P}$) 是鞅.

(b) 如果

$$\mathbf{E}\exp\left\{\frac{1}{2}\sum_{k=1}^{n}\left(\frac{\mu_k}{\sigma_k}\right)^2\right\} < \infty \quad \text{(即有诺维可夫条件成立)}$$

及

$$Z_\infty = \exp\left\{-\sum_{k=1}^{\infty}\frac{\mu_k}{\sigma_k}\xi_k - \frac{1}{2}\sum_{k=1}^{\infty}\left(\frac{\mu_k}{\sigma_k}\right)^2\right\},$$

则 $Z_n = (Z_n)_{n\geqslant 1}$ 是一致可积鞅, 概率为 1 地存在 $Z_\infty = \lim Z_n$, 且 $Z_n = \mathbf{E}(Z_\infty \,|\, \mathscr{F}_n)$, $n \geqslant 1$.

9. 设 $\mathscr{F} = \sigma(\bigcup \mathscr{F}_n)$. 采用上题的记号, 令

$$\widetilde{\mathbf{P}}(d\omega) = Z_\infty \mathbf{P}(d\omega).$$

证明, 如果 $\mathbf{E}\, Z_\infty = 1$, 则条件分布

$$\mathrm{Law}(h_n \,|\, \mathscr{F}_{n-1}; \widetilde{\mathbf{P}}) = \mathcal{N}(0, \sigma_n^2).$$

如果进一步, $\sigma_n^2 = \sigma_n^2(\omega)$ 与 ω 无关, 则

$$\mathrm{Law}(h_n \,|\, \widetilde{\mathbf{P}}) = \mathcal{N}(0, \sigma_n^2),$$

在此, 随机变量 $h_1, h_2, \cdots$ 是 (关于 $\widetilde{\mathbf{P}}$) 相互独立的.

10. 设 $X_n = e^{H_n}$, 其中 $H_n = h_1 + \cdots + h_n$, $n \geqslant 1$, 而 $h_k = \mu_k + \sigma_k \xi_k$, 又 μ_k, σ_k, ξ_k, $k \geqslant 1$, 均同第 8 题.

证明, 如果 ($\mathbf{P}$-a.s. 地) 有

$$\mu_k + \frac{\sigma_k^2}{2} = 0, \quad k \geqslant 1,$$

则 $X = (X_n, \mathscr{F}_n)_{n\geqslant 1}$ 是鞅.

假设上述条件对某些 $k \geqslant 1$ 不成立. 令

$$Z_\infty = \exp\left\{-\sum_{k=1}^{\infty}\left(\frac{\mu_k}{\sigma_k} + \frac{\sigma_k}{2}\right)\xi_k - \frac{1}{2}\sum_{k=1}^{\infty}\left(\frac{\mu_k}{\sigma_k} + \frac{\sigma_k}{2}\right)^2\right\}.$$

假设 $\mathbf{E}\, Z_\infty = 1$, 引入测度

$$\widetilde{\mathbf{P}}(d\omega) = Z_\infty \mathbf{P}(d\omega),$$

并令 $\mathscr{F}=\sigma(\bigcup \mathscr{F}_n)$.

证明, 关于测度 $\widetilde{\mathbf{P}}$, 序列 $(X_n,\ \mathscr{F}_n)_{n\geqslant 1}$ 是鞅, 其中 $X_n=e^{H_n}$.

§8. 相依随机变量之和的中心极限定理

1. 设 $\xi_n=\eta_n+\zeta_n,\ n\geqslant 1$, 其中 $\eta_n\xrightarrow{\mathrm{d}}\eta$, 而 $\zeta_n\xrightarrow{\mathrm{d}}0$. 证明, $\xi_n\xrightarrow{\mathrm{d}}\eta$.

2. 设 $(\xi_n(\varepsilon)),\ n\geqslant 1,\ \varepsilon>0$, 为随机变量族, 对每个 $\varepsilon>0$, 当 $n\to\infty$ 时, 都有 $\xi_n(\varepsilon)\xrightarrow{\mathbf{P}}0$. 试利用诸如第二章第 10 节第 11 题中的命题, 证明, 存在数列 $\varepsilon_n\downarrow 0$, 使得 $\xi_n(\varepsilon_n)\xrightarrow{\mathbf{P}}0$.

提示: 应当选取 $\varepsilon_n\downarrow 0$, 使得

$$\mathbf{P}\{|\xi_n(\varepsilon_n)|\geqslant 2^n\}\leqslant 2^{-n},\quad n\geqslant 1.$$

3. 设 $(\alpha_k^n),\ 1\leqslant k\leqslant n,\ n\geqslant 1$, 是复值随机变量, ($\mathbf{P}$-a.s. 地) 有

$$\sum_{k=1}^{n}|\alpha_k^n|\leqslant C,\quad |\alpha_k^n|\leqslant a_n\downarrow 0.$$

证明, 当 $n\to\infty$ 时, ($\mathbf{P}$-a.s. 地) 有

$$\lim_n \prod_{k=1}^{n}(1+\alpha_k^n)e^{-\alpha_k^n}=1.$$

4. 试证明关于定理 1 的注 2 中的命题.

5. 试证明第 4 小节中所陈述的引理.

6. 试证明定理 3.

7. 试证明定理 5.

8. 设 $\xi=(\xi_n)_{-\infty<n<\infty}$ 是独立同分布的随机变量序列, 有 $\mathbf{E}\,\xi_n=0,\ \mathbf{D}\,\xi_n<\infty$. 定义序列 $\eta=(\eta_n)_{n\geqslant 1}$ 如下:

$$\eta_n=\sum_{j=-\infty}^{\infty}c_{n-j}\xi_j,\quad \sum_{j=-\infty}^{\infty}|c_j|^2<\infty.$$

假设

$$D_n^2=\mathbf{E}\,(\eta_1+\cdots+\eta_n)^2\to\infty.$$

证明,

$$\mathbf{P}\left\{\frac{\eta_1+\cdots+\eta_n}{D_n}\leqslant x\right\}\to\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-t^2/2}dt.$$

9. 设 $(\Omega^n,\ \mathscr{F}^n,\ (\mathscr{F}_k^n)_{0\leqslant k\leqslant n},\ \mathbf{P}^n)$ 是可滤概率空间, $n\geqslant 1$. 对每个 $n\geqslant 1$, 都给定了一组随机变量 $\xi^n=(\xi_k^n)_{1\leqslant k\leqslant n}$, 其中 ξ_k^n 是 $\mathscr{F}_k^n$ 可测的.

设 μ 是 $(\mathbb{R},\ \mathscr{B}(\mathbb{R}))$ 上的无穷可分分布, 具有特征函数 (b,c,F) (参阅第三章第 6 节第 17 题, 其中 $h=h(x)$ 是连续的截断函数.)

证明, 欲随机变量 $Z^n=\sum_{k=1}^{n}\xi_k^n$ 的分布弱收敛到 μ, 只需如下各条件成立:

$$\sup_{1\leqslant k\leqslant n}\mathbf{P}^n\{|\xi_k^n|>\varepsilon\,|\,\mathscr{F}_{k-1}^n\}\xrightarrow{\mathbf{P}}0,\quad \varepsilon>0,$$

$$\sum_{1\leqslant k\leqslant n}\mathbf{E}^n\left[h(\xi_k^n)\,|\,\mathscr{F}_{k-1}^n\right]\xrightarrow{\mathbf{P}}b,$$

$$\sum_{1\leqslant k\leqslant n}\left(\mathbf{E}^n\left[h^2(\xi_k^n)\,|\,\mathscr{F}_{k-1}^n\right]-(\mathbf{E}^n\left[h(\xi_k^n)\,|\,\mathscr{F}_{k-1}^n\right])^2\right)\xrightarrow{\mathbf{P}}\widetilde{c},$$

$$\sum_{1\leqslant k\leqslant n}\mathbf{E}^n\left[g(\xi_k^n)\,|\,\mathscr{F}_{k-1}^n\right]\xrightarrow{\mathbf{P}}F(g),\quad g\in\mathfrak{G}_1,$$

其中, $\widetilde{c}=c+\int h^2(x)F(dx)$, 而 $\mathfrak{G}_1=\{g\}$ 是所有形如 $g_a(x)=(a|x|-1)^+\wedge 1$ 的函数的集合, 其中 a 是有理数, 又 $F(g)=\int g(x)F(dx)$.

10. 设 $\xi_0,\xi_1,\xi_2,\cdots$ 是严平稳随机变量序列, 有 $\mathbf{E}\,\xi_0=0$. 令 (参阅第六章第 3 节第 5 题)

$$\alpha_k=\sup\ |\mathbf{P}(A\cap B)-\mathbf{P}(A)\mathbf{P}(B)|,\quad k\geqslant 1,$$

其中的上确界系对所有满足如下条件的集合所取:

$$A\in\mathscr{F}_0=\sigma\,(\xi_0),\quad B\in\mathscr{F}_k^{\infty}=\sigma\,(\xi_k,\xi_{k+1},\cdots).$$

令

$$X_t^n=\frac{1}{\sqrt{n}}\sum_{k=1}^{[nt]}\xi_k,\quad t\geqslant 0.$$

证明, 如果对某个 $p>2$, 强混合系数 $\{\alpha_k,\ k\geqslant 1\}$ 满足条件

$$\sum_{k\geqslant 1}\alpha_k^{(p-2)/p}<\infty,$$

且对该 p, 有 $\mathbf{E}\,|\xi_0|^p<\infty$, 则随机变量 $X_{t_1}^n,\cdots,X_{t_k}^n$ 的联合分布 $P_{t_1,\cdots,t_k}^n$ 弱收敛到随机向量 $(\sqrt{c}B_{t_1},\cdots,\sqrt{c}B_{t_k})$ 的分布 $P_{t_1,\cdots,t_k}$, 其中 $B=(B_t)_{t\geqslant 0}$ 是布朗桥, 而 c 是如下形式的常数:

$$c=\mathbf{E}\,\xi_0^2+2\sum_{k\geqslant 1}\mathbf{E}\,\xi_k^2.$$

§9. 伊藤公式的离散版本

1. 试证明 (15) 式.

2. 在关于随机游动 $S=(S_n)_{n\geqslant 0}$ 的中心极限定理的基础上, 证明

$$\mathbf{E}\,|S_n| \sim \sqrt{\frac{2}{\pi}n}, \quad n\to\infty.$$

(试比较第一章第 9 节第 3 题中的提示.)

注: 在第七章第 9 节例 2 中的 (17) 式和 (18) 式中的 $1/2\pi$ 应当为 $2/\pi$.

3. 试证明 (22) 式.

4. (24) 式可对所有函数 $F\in C^2$ 成立, 试证之.

5. 试将 (11) 式推广到非齐次 d 维向量的场合, 其中以分布函数 $F(k,X_k^1,\cdots,X_k^d)$ 取代 $F(X_k)$.

6. 设 $f(x)=F'(x)$. 我们来考察等式

$$F(X_n)=F(X_0)+\sum_{k=1}^n f(X_{k-1})\Delta X_k+\sum_{k=1}^n (F(X_k)-F(X_{k-1})-f(X_{k-1})\Delta X_k),$$

尽管该式其貌不扬, 它却可以视为离散时间场合下的一个换元公式 (伊藤公式的离散版本).

试给出一个映射, 以表明如何可以由该式得到关于布朗运动场合下的二次连续可微函数 $F=F(x)$ 的伊藤换元公式 ((24) 式).

7. 试将上题中的公式推广到非齐次 d 维向量的场合 (以 $F(k,X_k^1,\cdots,X_k^d)$ 取代 $F(X_k)$).

8. (续角谷公式的离散版本; 参阅第一章第 9 节第 3 题) 考察对称伯努利系统 (即: 独立同分布的随机变量序列 $\xi_1,\xi_2,\cdots$, 有 $\mathbf{P}\{\xi_n=1\}=\mathbf{P}\{\xi_n=-1\}=1/2$, $n\geqslant 1$), 令 $S_0=0$, $S_n=\xi_1+\cdots+\xi_n$, $n\geqslant 1$. 对 $x\in\mathbb{Z}=\{0,\pm1,\pm2,\cdots\}$, 记

$$N_n(x)=\#\{k,\ 0\leqslant k\leqslant n:\ S_k=x\},$$

亦即满足等式 $S_k=x$ 和 $0\leqslant k\leqslant n$ 的整数 k 的个数.

证明角谷公式的如下离散版本

$$|S_n-x|=|x|+\sum_{k=1}^n \operatorname{sign}(S_{k-1}-x)\Delta S_k+N_n(x).$$

注: 如果 $B=(B_t)_{t\geqslant 0}$ 是布朗运动, 则由著名的角谷公式知

$$|B_t-x|=|x|+\int_0^t \operatorname{sign}(B_s-x)dB_s+N_t(x), \quad t\geqslant 0,$$

其中 $N_t(x)$ 是区间 $[0,t]$ 中布朗运动处于水平 $x\in\mathbb{R}$ 的局部时. (莱维关于局部时 $N_t(x)$ 的定义是:

$$N_t(x)=\lim_{\varepsilon\downarrow 0}\frac{1}{2\varepsilon}\int_0^t I(|B_s-x|\leqslant\varepsilon)dx,$$

可参阅, 例如, 文献 [12], [97].)

§10. 保险中破产概率的计算. 鞅方法

1. 证明, $N=(N_t)_{t\geqslant 0}$ (在假设 A 之下, 参阅《概率》第二卷 p201) 是独立增量过程.

2. 证明, $X=(X_t)_{t\geqslant 0}$ 也是独立增量过程.

3. 考察克拉默–林德伯格模型, 并对 σ_i, $i=1,2,\cdots$ 为相互独立的几何分布随机变量 (即 $\mathbf{P}\{\sigma_i=k\}=q^{k-1}p$, $k\geqslant 1$) 的情形陈述相应的定理.

4. 设 $N=(N_t)_{t\geqslant 0}$ 为由 (3) 式所定义的 (以 λ 为参数的) 泊松过程. 令 $0=t_0<t_1<t_2<\cdots<t_n$, $0\leqslant k_1\leqslant k_2\leqslant\cdots\leqslant k_n$. 证明如下的 "马氏性":

$$\mathbf{P}\{N_{t_k}=k_n \mid N_{t_1}=k_1,\cdots,N_{t_{n-1}}=k_{n-1}\}=\mathbf{P}\{N_{t_k}=k_n \mid N_{t_{n-1}}=k_{n-1}\}.$$

5. 设 $N=(N_t)_{t\geqslant 0}$ 是标准 (即参数为 $\lambda=1$ 的) 泊松过程, 而 $\lambda(t)$ 是非降右连续函数, 有 $\lambda(0)=0$. 考察过程 $N\circ\lambda=(N_{\lambda(t)})_{t\geqslant 0}$, 并研究其性质 (有限维分布, 矩等等).

6. 设 $(T_1,\cdots,T_n)$ 是泊松过程的前 n 个跳跃时刻, 而 $(X_1,\cdots,X_n)$ 是相互独立的区间 $[0,t]$ 上的均匀分布随机变量, $X_1^{(n)},\cdots,X_n^{(n)}$ 是 $X_1,\cdots,X_n$ 的秩序统计量. 证明,

$$\text{Law}\,(T_1,\cdots,T_n \mid N_t=n)=\text{Law}\,(X_1^{(n)},\cdots,X_n^{(n)}),$$

即随机向量 $(T_1,\cdots,T_n)$ 在条件 $N_t=n$ 之下的条件分布与随机向量 $(X_1^{(n)},\cdots,X_n^{(n)})$ 的分布相同.

7. 证明, 对于泊松过程, 当 $s<t$ 时, 有

$$\mathbf{P}(N_s=m \mid N_t=n)=\begin{cases} C_n^m(s/t)^m(1-s/t)^{n-m}, & m\leqslant n,\\ 0, & m>n.\end{cases}$$

8. 在初等概率论中就已经证明了: 如果 X_1 与 X_2 是两个相互独立的随机变量, 分别服从参数为 λ_1 与 λ_2 的泊松分布, 则 X_1+X_2 是服从参数为 $\lambda_1+\lambda_2$ 的泊松分布的随机变量. 试证明其逆命题 (拉伊可夫): 设 X_1 与 X_2 是两个相互独立的随机变量, 使得 X_1+X_2 服从泊松分布, 则 X_1 与 X_2 均服从泊松分布.

9. 设 $N=(N_t)_{t\geqslant 0}$ 是标准泊松过程, θ 是与 N 独立的正值随机变量. 我们来考察干扰过程 $\widetilde{N}=(\widetilde{N}_t)_{t\geqslant 0}$, 其中 $\widetilde{N}_t=N_{t\theta}$. 证明这一过程的如下各种极限性状:

(a) 强大数定律:

$$\frac{\widetilde{N}_t}{t}\to\theta\quad(\mathbf{P}\text{-a.s.}),\quad t\to\infty$$

(参阅第四章第 3 节第 4 小节例 4).

(b) 中心极限定理:

$$\mathbf{P}\left\{\frac{\widetilde{N}_t-\theta t}{\sqrt{\theta t}}\leqslant x\right\}\to\Phi(x),\quad t\to\infty.$$

(c) 若 $0<\mathbf{D}\,\theta<\infty$, 则,

$$\frac{\widetilde{N}_t-\mathbf{E}\widetilde{N}_t}{\sqrt{\mathbf{D}\widetilde{N}_t}}\to\frac{\theta-\mathbf{E}\theta}{\sqrt{\mathbf{D}\theta}}.^{①}$$

10. 证明, 等于给定的 $u>0$, "破产函数"

$$\psi(u)=\mathbf{P}\left\{\inf_{t\geqslant 0}X_t\leqslant 0\right\}\quad(=\mathbf{P}\{T<\infty\})$$

可以表示为

$$\psi(u)=\mathbf{P}\left\{\sup_{n\geqslant 1}Y_n\geqslant u\right\},$$

其中 $Y_n=\sum\limits_{i=1}^{n}(\xi_i-c\sigma_i)$.

再证明, 第 10 节中用鞅方法所得到的估计式 $\psi(u)\leqslant e^{-Ru}$ 可以用更为初等的方法推出. 意即: 如果记 $\psi_n(u)=\mathbf{P}\left\{\max_{1\leqslant k\leqslant n}Y_k\geqslant u\right\}$, $n\geqslant 1$, 则显然有 $\psi_1(u)\leqslant e^{-Ru}$. 余下再用数学归纳法即可得知, 对一切 $n>1$, 都有 $\psi_n(u)\leqslant e^{-Ru}$.

11. 破产时刻的定义是 $T=\inf\{t\geqslant 0:\ X_t\leqslant 0\}$, 但有时也以 $\widetilde{T}=\inf\{t\geqslant 0:\ X_t<0\}$ 作为破产时刻. 试分析, 如果将 T 换为 $\widetilde{T}$, 那么第 10 节中的各种结果将会怎样变化?

12. 作为第 2 小节所引入的 (齐次) 泊松过程的推广, 我们来考察非齐次泊松过程 $N=(N_t)_{t\geqslant 0}$, 其定义为

$$N_t=\sum_{i\geqslant 1}I(T_i\leqslant t),$$

其中 $T_i=\sigma_1+\cdots+\sigma_i$, 随机变量 σ_i 独立同分布, 服从如下分布:

$$\mathbf{P}\{\sigma_i\leqslant t\}=1-\exp\left\{-\int_0^t\lambda(s)ds\right\},$$

①原书为: $\frac{\widetilde{N}_t-\mathbf{E}\widetilde{N}_t}{\sqrt{\mathbf{D}N_t}}\to\frac{\theta-\mathbf{E}\theta}{\sqrt{\mathbf{D}\theta}}$ —— 译者注.

在这里, $\lambda(t)$ 称为强度函数, 有 $\lambda(t)\geqslant 0$, $\int_0^t\lambda(s)ds<\infty$ 和 $\int_0^\infty\lambda(s)ds=\infty$. 证明,

$$\mathbf{P}\{N_t<k\}=\mathbf{P}\{T_k>t\}=\sum_{i=0}^{k-1}\exp\left\{-\int_0^t\lambda(s)ds\right\}\frac{\left(\int_0^t\lambda(s)ds\right)^i}{i!}.$$

13. 设 $N=(N_t)_{t\geqslant 0}$ 是非齐次泊松过程 (参阅第 12 题), $(\xi_n)_{n\geqslant 0}$ 是独立同分布的随机变量序列, 且与过程 N 独立. 设 $g=g(t,x)$ 是 $\mathbb{R}\times\mathbb{R}$ 上的非负函数. 证明如下的康贝尔公式

$$\mathbf{E}\sum_{n=1}^{\infty}g(T_n,\,\xi_n)I(T_n\leqslant t)=\int_0^t\mathbf{E}\,[g(s,\,\xi_1)]\lambda(s)ds.$$

14. 设 $N=(N_t)_{t\geqslant 0}$ 是齐次泊松过程, $N_0=0$, $N_t=\sum\limits_n I(T_n\leqslant t)$, $t>0$, 而随机变量 $\sigma_{n+1}=T_{n+1}-T_n$ $(n\geqslant 0,\ T_0=0)$ 独立同分布, 有

$$\mathbf{P}\{\sigma_{n+1}\geqslant x\}=e^{-\lambda x},\ x\geqslant 0.$$

令 $U_t=t-T_{N_t}$, $V_t=T_{N_t+1}$. 证明,

$$\mathbf{P}\{U_t\leqslant u,\ V_t\leqslant v\}=[I_{\{u\geqslant t\}}+I_{\{u<t\}}(1-e^{-\lambda u})](1-e^{-\lambda\nu}).$$

(特别地, 由此可以得知: 对于每个 $t>0$, 随机变量 U_t 与 V_t 都相互独立, 并且 V_t 具有参数为 λ 的泊松分布.) 试求概率 $\mathbf{P}\{T_{N_t+1}-T_{N_t}\geqslant x\}$. 证明, $\mathbf{P}\{T_{N_t+1}-T_{N_t}\geqslant x\}\neq e^{-\lambda x}$ $(=\mathbf{P}\{T_{n+1}-T_n\geqslant x\})$. 并且证明, 当 $t\to\infty$ 时, $T_{N_t+1}-T_{N_t}$ 的分布弱收敛到两个相互独立的均服从参数为 λ 的指数分布的随机变量之和的分布.

§11. 随机金融数学的基本定理. 无套利的鞅特征

1. 证明, 当 $N=1$ 时, 无套利条件等价于不等式 (18) 成立 (假定 $\mathbf{P}\{\Delta S_1=0\}<1$).

2. 证明, 在 (第 4 小节) 引理 4 的证明中用到了条件 (19).

3. 证明, (第 5 小节) 例 1 中的测度 $\widetilde{\mathbf{P}}$ 是鞅测度, 而且这在族 $\mathbf{M}(\mathbf{P})$ 中是唯一的.

4. 试研究 (第 5 小节) 例 2 中所构造的鞅测度的唯一性.

5. 证明, 在 $(B,\,S)$ 模型中, 假设 "$|\mathbf{M}(\mathbf{P})|=1$" 中包含了关于 $\frac{S_n}{B_n}$ $(1\leqslant n\leqslant N)$ 的 "两点条件" (第 6 小节).

6. 根据关于定理 1 的注 1, 如果 $N<\infty$ 且 $d<\infty$, 则有 "第一基本定理" 成立. 试举例说明, 如果 $d=\infty$, 则尽管无套利存在, 鞅测度也未必存在.

7. 作为定义 1 的补充中, 我们说 $(B,\,S)$-市场在弱意义下是无套利的, 如果对一切 "自融资组合" $\pi=(\beta,\gamma)$, 其中 $X_0^\pi=0$, 而 $X_n^\pi\geqslant 0$ ($\mathbf{P}$-a.s.), $n\leqslant N$, 我们都有

$X_N^\pi = 0$ ($\mathbf{P}$-a.s.). 说 (B, S)-市场在强意义下是无套利的, 如果对一切 "自融资组合" π, 其中 $X_0^\pi = 0$, 而 $X_N^\pi \geqslant 0$ ($\mathbf{P}$-a.s.), 我们都有 $X_n^\pi = 0$ ($\mathbf{P}$-a.s.), $0 \leqslant n \leqslant N$.

设定理 1 中的假设成立. 证明, 如下各条件相互等价:

(i) (B, S)-市场是无套利的.

(ii) (B, S)-市场在弱意义下是无套利的.

(iii) (B, S)-市场在强意义下是无套利的.

8. (如同在定理 1 中), 设

$$\mathbf{M}(\mathbf{P}) = \left\{\widetilde{\mathbf{P}} \sim \mathbf{P}: \ \frac{S}{B} \text{ 是 } \widetilde{\mathbf{P}} \text{ 鞅}\right\}$$

是所有鞅测度的集合,

$$\mathbf{M}_{\mathrm{loc}}(\mathbf{P}) = \left\{\widetilde{\mathbf{P}} \sim \mathbf{P}: \ \frac{S}{B} \text{ 是 } \widetilde{\mathbf{P}} \text{ 局部鞅}\right\};$$

$$\mathbf{M}_{\mathrm{b}}(\mathbf{P}) = \left\{\widetilde{\mathbf{P}} \sim \mathbf{P}: \ \widetilde{\mathbf{P}} \in \mathbf{M}(\mathbf{P}), \text{且对某个常数 } C(\widetilde{\mathbf{P}}), \text{有 } (\mathbf{P}\text{-a.s.}) \text{ 密度 } \frac{d\widetilde{\mathbf{P}}}{d\mathbf{P}}(\omega) \leqslant C(\widetilde{\mathbf{P}})\right\}.$$

证明, 在定理 1 的条件下, 如下各条件相互等价:

$$\text{(i) } \mathbf{M}(\mathbf{P}) \neq \varnothing, \qquad \text{(ii) } \mathbf{M}_{\mathrm{loc}}(\mathbf{P}) \neq \varnothing, \qquad \text{(iii) } \mathbf{M}_{\mathrm{b}}(\mathbf{P}) \neq \varnothing.$$

§12. 无套利模型中与 "对冲" 有关的核算

1. 试对第 11 节第 5 小节例 2 中所考察的 (B, S)-市场模型, 求出带有 $f_N = (S_N - K)^+$ 的标准买入期权的价格 $\mathbb{C}(f_N; \mathbf{P})$.

2. 请尝试证明 (10) 式的反向不等式.

3. 试证明 (12) 式, 并请尝试证明 (13) 式.

4. 试给出 (23) 式的详细结论.

5. 试证明 (25) 式与 (28) 式.

6. 试给出 (32) 式的详细结论.

7. 在一步 CRR-模型

$$B_1 = B_0(1+r), \quad S_1 = S_0(1+\rho)$$

中 (参阅第 7 小节 (17) 式), 假定了 ρ 只取两个值 a 与 b, 且 $-1 < a < r < b$. 我们现在假定 ρ 服从区间 $[a, b]$ 上的均匀分布.

证明, 在这种情况下, 上估计 ($S_0 =$ 常数)

$$\widehat{\mathbb{C}}(f; \mathbf{P}) = \inf\{x: \ \exists \pi \text{ 有 } X_0^\pi = x \text{ 与 } X_1^\pi \geqslant f(S_0(1+\rho)), \ \forall \rho \in [a, b]\},$$

其中 $f(S_0(1+\rho))$ 是下凸的连续函数, $\rho \in [a, b]$, 重合于 CRR-模型 (其中 $\mathbf{P}\{\rho = b\} = p$, $\mathbf{P}\{\rho = a\} = 1-p$, $0 < p < 1$) 的上估计 (参阅 (19) 式, 在其中令 $N = 1$), 因而有

$$\widehat{\mathbb{C}}(f; \mathbf{P}) = \frac{r-a}{b-a} \cdot \frac{f(S_0(1+b))}{1+r} + \frac{b-r}{b-a} \cdot \frac{f(S_0(1+a))}{1+r}.$$

8. (关于布莱克–舒尔斯公式) 作为离散 (时间) (B, S)-市场, 其中 $B = (B_n)_{0\leqslant n\leqslant N}$, $S = (S_n)_{0\leqslant n\leqslant N}$ (参阅《概率》第二卷 p219 的推广, 我们来考察连续 (时间) (B, S)-市场, 其中 $B = (B_t)_{0\leqslant t\leqslant T}$, $S = (S_t)_{0\leqslant t\leqslant T}$, 且

$$B_t = B_0 e^{-rt}, \quad r \geqslant 0, \quad B_0 > 0$$

和

$$S_t = S_0 e^{\mu t + \sigma W_t}, \tag{$*$}$$

其中 $S_0 > 0$, 而 $W = (W_t)_{0\leqslant t\leqslant T}$ 是维纳过程 (布朗运动).

与买入期权支付函数 $f_N = (S_N - K)^+$(参阅《概率》第二卷 p219 中的 (1) 式) 中类似, 我们也认为在 $(*)$ 式中, 有

$$\mu = r - \frac{\sigma^2}{2}.$$

试在该假设之下, 证明,

(a) 序列 $(\frac{S_t}{B_t})_{0\leqslant t\leqslant T}$ 是鞅.

(b) 由

$$\mathbb{C}(f_T; \mathbf{P}) = B_0 \mathbf{E} \frac{f_T}{B_T}$$

所定义的相应于买入期权及顾客等的 $\mathbb{C}(f_T; \mathbf{P})$ 满足布莱克–舒尔斯公式:

$$\mathbb{C}(f_T; \mathbf{P}) = S_0 \Phi\left(\frac{\ln\frac{S_0}{k} + T\left(r + \frac{\sigma^2}{2}\right)}{\sigma\sqrt{T}}\right) - K e^{-rT} \Phi\left(\frac{\ln\frac{S_0}{k} + T\left(r - \frac{\sigma^2}{2}\right)}{\sigma\sqrt{T}}\right).$$

提示: 对于 (b), 应当先证 $\mathbb{C}(f_T; \mathbf{P}) = e^{-rT}(a e^{b\xi - b^2/2} - K)^+$, 其中 $a = S_0 e^{rT}$, $b = \sigma\sqrt{T}$, 而 $\xi \sim \mathcal{N}(0, 1)$. 然后通过计算直接证明

$$\mathbf{E}\left(a e^{b\xi - \frac{b^2}{2}} - K\right)^+ = a\Phi\left(\frac{\ln\frac{a}{k} + \frac{1}{2}b^2}{b}\right) - K\Phi\left(\frac{\ln\frac{a}{k} - \frac{1}{2}b^2}{b}\right).$$

§13. 最优停止问题. 鞅方法

1. 证明, 在第 3 小节引理的证明中所构造的随机变量 $\xi(\omega) = \sup\limits_{\alpha\in\mathfrak{U}_0} \xi_\alpha(\omega)$ 满足关于上确界存在性定义中的要求 a) 与 b).

提示: 在 $\alpha \notin \mathfrak{U}_0$ 的情况下, 考察 $\mathbf{E}\max(\xi(\omega),\ \xi_\alpha(\omega))$.

2. 证明, 随机变量 $\xi(\omega) = \tan\widetilde{\xi}(\omega)$(参阅第 3 小节引理证明的末尾) 亦满足要求 a) 与 b).

3. 设 $\xi_1, \xi_2, \cdots$ 是独立同分布的随机变量序列, 有 $\mathbf{E}|\xi_1| < \infty$. 我们来考察关于 ($\mathfrak{M}_1^\infty = \{\tau:\ 1 \leqslant \tau < \infty\}$ 类中的) 最佳停时问题:

$$V^* = \sup_{\tau\in\mathfrak{M}_1^\infty} \mathbf{E}\left(\max_{i\leqslant\tau} \xi_i - c\tau\right).$$

令 $\tau^* = \inf\{n \geqslant 1:\ \xi_n \geqslant A^*\}$, 其中 A^* 是方程 $\mathbf{E}(\xi_i - A^*) = c$ 的唯一解 ($\inf\varnothing = \infty$). 证明, 如果 $\mathbf{P}\{\tau^* < \infty\} = 1$, 则 τ^* 是使得 $\mathbf{E}\left(\max\limits_{i\leqslant\tau} \xi_i - c\tau\right)$ 存在的所有停时 τ 的类中的最佳停时.

再证明 $V^* = A^*$.

4. 在本题和下一题中, 采用如下记号:

$$\mathfrak{M}_n^\infty = \{\tau:\ n \leqslant \tau < \infty\}, \qquad V_n^\infty = \sup_{\tau\in\mathfrak{M}_n^\infty} \mathbf{E} f_\tau,$$

$$v_n^\infty = \operatorname*{ess\,sup}_{\tau\in\mathfrak{M}_n^\infty} \mathbf{E}(f_\tau|\mathscr{F}_n), \qquad \tau_n^\infty = \inf\{k \geqslant n:\ v_n^\infty = f_n\}.$$

假设

$$\mathbf{E}\sup f_n^- < \infty,$$

证明, 对于极限随机变量[①]

$$\tilde{v}_n = \lim_{N\to\infty} v_n^N,$$

有如下各断言成立:

(a) 对于一切 $\tau \in \mathfrak{M}_n^\infty$, 都有

$$\tilde{v}_n \geqslant \mathbf{E}(f_\tau|\mathscr{F}_n).$$

(b) 如果 $\tau_n^\infty \in \mathfrak{M}_n^\infty$, 则有

$$\tilde{v}_n = \mathbf{E}(f_{\tau_n^\infty}|\mathscr{F}_n),$$

$$\tilde{v}_n = v_n^\infty \quad (= \operatorname*{ess\,sup}_{\tau\in\mathfrak{M}_n^\infty} \mathbf{E}(f_\tau|\mathscr{F}_n)).$$

[①] 随机变量 v_n^N 的定义参阅《概率》第二卷第七章第 13 节 (6) 式 —— 译者注.

5. 设 $\tau_n^\infty \in \mathfrak{M}_n^\infty$. 试由上题中的断言 (a) 与 (b) 推出, τ_n^∞ 在如下意义下是最佳停时:

$$\operatorname*{ess\,sup}_{\tau\in\mathfrak{M}_n^\infty} \mathbf{E}(f_\tau|\mathscr{F}_n) = \mathbf{E}(f_{\tau_n^\infty}|\mathscr{F}_n) \quad (\mathbf{P}\text{-a.s.})$$

及

$$\sup_{\tau\in\mathfrak{M}_n^\infty} \mathbf{E} f_\tau = \mathbf{E} f_{\tau_n^\infty},$$

亦即 $V_n^\infty = \mathbf{E} f_{\tau_n^\infty}$.

6. 设 $(\Omega, \mathscr{F}, \mathbf{P})$ 为概率空间, 而 $\Sigma = \{\xi_\alpha(\omega);\ \alpha \in \mathfrak{U}\}$ 是满足条件: "对任何 $\alpha \in \mathfrak{U}$, 都有 $\mathbf{E}|\xi_\alpha| \leqslant C$" 的随机变量族. 假设 Σ 足够 "丰富", 使得只要 $\alpha_1, \alpha_2 \in \mathfrak{U}$ 与 $A \in \mathscr{F}$, 则对任何 $\xi_{\alpha_1} \in \Sigma$, $\xi_{\alpha_2} \in \Sigma$, 都有

$$\xi = \xi_{\alpha_1} I_A + \xi_{\alpha_2} I_{\bar{A}} \in \Sigma.$$

(此时我们说族 Σ 是 "可以缝补" 的.) 对这样的族 Σ, 令

$$\mathbf{Q}(A) = \sup_{\alpha\in\mathfrak{U}} \mathbf{E}\xi_\alpha I_A, \quad A \in \mathscr{F}.$$

证明,

(a) $\mathbf{Q} = \mathbf{Q}(\cdot)$ 是 σ-可加的集合函数.

(b) $\mathbf{Q} \ll \mathbf{P}$.

(c) 拉东–尼科迪姆导数

$$\frac{d\mathbf{Q}}{d\mathbf{P}} = \operatorname*{ess\,sup}_{\alpha\in\mathfrak{U}} \quad (\mathbf{P}\text{-a.s.}).$$

(断言 (c) 可以视为对于 "可以缝补" 的随机变量族的本质上确界的存在性的证明.)

证明, 如果将条件 "$\mathbf{E}|\xi_\alpha| \leqslant C,\ \alpha \in \mathfrak{U}$" 换为 "$\mathbf{E}\,\xi_\alpha^- < \infty,\ \alpha \in \mathfrak{U}$", 则断言 (a), (b), (c) 仍可成立.

7. 设 $\mathfrak{M}_n^\infty = \{\tau:\ n \leqslant \tau < \infty\}$. 证明, 如果 $\tau_1, \tau_2 \in \mathfrak{M}_n^\infty$, 而 $A \in \mathscr{F}_n$, 则有 $\tau = \tau_1 I_A + \tau_2 I_{\bar{A}}$ 属于 $\mathfrak{M}_n^\infty$.

8. 设 $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geqslant 0}, \mathbf{P})$ 为可滤概率空间, f_n 为 $\mathscr{F}_n$ 可测的函数, 有 $\mathbf{E} f_n^- < \infty$, $n \geqslant 0$. 证明, 随机变量族 $\{\mathbf{E}(f_\tau|\mathscr{F}_n);\ \tau \in \mathfrak{M}_n^\infty\}$ (对每个固定的 $n \geqslant 0$) 是 "可以缝补" 的.

第八章　形成马尔可夫链的随机变量序列

§1. 定义和基本性质

1. 试将关于定理 1 的证明中所建立的断言的证明分别作为本题的 a, b, c 小题 (参阅第 3 小节).

2. 证明, 定理 2 中的函数 $P_{n+1}(B-X_n(\omega))$ 关于 ω 是 $\mathscr{F}_n$ 可测的.

3. 试由第二章第 2 节引理 3 中的断言推出性质 (11) 和 (12).

4. 证明性质 (20) 和 (27).

5. 证明关系式 (33).

6. 证明第 8 小节末尾所陈述的断言 (i),(ii) 和 (iii).

7. 可否由马尔可夫性质 (3) 推出如下各性质:

$$\mathbf{P}\{X_{n+1}\in B\mid X_0\in B_0,\,X_1\in B_1,\cdots,X_n\in B_n\}=\mathbf{P}\{X_{n+1}\in B\mid X_n\in B_n\},$$

其中, $B_0,B_1,\cdots,B_n$ 和 B 是 $\mathscr{E}$ 中的集合, 而 $\mathbf{P}\{X_0\in B_0,\,X_1\in B_1,\cdots,X_n\in B_n\}>0$?

8. 我们来观察一段圆柱状粉笔, 其长度为 1. "随机地" 将它折为两段. 再 "随机地" 将左半段折为两段, 并如此一直下去. 将折了 n 次之后的左段的长度记为 X_n, 令 $\mathscr{F}_n=\sigma(X_1,\cdots,X_n)$. 则有 $X_0=1$, 而当 $X_n=x$ 时, X_{n+1} 的条件分布是区间 $[0,x]$ 上的均匀分布.

　　证明, 如此所得到的序列 $(X_n)_{n\geqslant 0}$ 是齐次马尔可夫链, 并且对任何 $\alpha>-1$, 序列

$$M_n=(1+\alpha)X_n^{\alpha},\quad n\geqslant 0$$

是非负鞅. 证明, 概率为 1 地对任何 $0<p<e$, 都有

$$\lim_n p^nX_n=0,$$

而对于 $p>e$, 则有

$$\lim_n p^nX_n=\infty.$$

(在随机折断中, 平均来说, 应当是对半分, 由此通常会认为 X_n 趋于 0 的速度应为 2^{-n}, 然而 $\lim_n p^nX_n=0$ ($\mathbf{P}$-a.s.) 的结果却表明 X_n 趋于 0 的速度要快得多, 几乎达到 e^{-n} 阶.)

9. 设 $\xi_1,\xi_2,\cdots$ 为独立同分布的随机变量序列, 有 $\mathbf{P}\{\xi_n=1\}=\mathbf{P}\{\xi_n=-1\}=1/2$, $n\geqslant 1$(伯努利序列). 令 $S_0=0$, $S_n=\xi_1+\cdots+\xi_n$, 而 $M_n=\max\{S_k:\ 0\leqslant k\leqslant n\}$, $n\geqslant 1$.

　　(a) 试问: 序列 $(|S_n|)_{n\geqslant 0}$, $(|M_n|)_{n\geqslant 0}$ 与 $(M_n-S_n)_{n\geqslant 0}$ 是否构成马尔可夫链?

　　(b) 如果 $S_0=x\neq 0$, $S_n=x+\xi_1+\cdots+\xi_n$, 那么上述三个序列是否构成马尔可夫链?

10. 设 $X=(X_n)_{n\geqslant 0}$ 为马尔可夫链, 相位空间为 $E=\{-1,0,1\}$. 设 $p_{i,j}>0$, $i,j\in E$, 试给出序列 $(|X_n|)_{n\geqslant 0}$ 构成马尔可夫链的充分必要条件.

11. 试举出这样的例子: 序列 $X=(X_n)_{n\geqslant 0}$ 不是马尔可夫链, 但是却满足柯尔莫戈洛夫–查普曼方程.

12. 设 $X=(X_n)_{n\geqslant 0}$ 是 (广义) 马尔可夫链, 而 $Y_n=X_{n+1}-X_n$, $n\geqslant 0$. 证明, $(X,Y)=((X_n)_{n\geqslant 0},\ (Y_n)_{n\geqslant 0})$ 仍然是马尔可夫链. 试问, 如下各序列是否是马尔可夫链: $(X_n,\,X_{n+1})_{n\geqslant 0}$, $(X_{2n})_{n\geqslant 0}$ 以及 $X=(X_{n+k})_{n\geqslant 0}$, 其中 $k\geqslant 1$?

13. 设 X_n, $n\geqslant 0$, 是在可数空间 E 中取值的随机变量. 称序列 $X=(X_n)_{n\geqslant 0}$ 构成$r\geqslant 1$ 阶的马尔可夫链, 如果对一切 $i_0,\cdots,i_{n+1}$, $n\geqslant r$ 都有

$$\begin{aligned}\mathbf{P}\{X_{n+1}=i_{n+1}|X_0=i_0,\cdots,X_n=i_n\}\\=\mathbf{P}\{X_{n+1}=i_{n+1}|X_{n-r+1}=i_{n-r+1},\cdots,X_n=i_n\}.\end{aligned}$$

　　令 $\widetilde{X}_n=(X_n,X_{n+1},\cdots,X_{n+r-1})$, $n\geqslant 0$. 证明, 序列 $\widetilde{X}=(\widetilde{X}_n)_{n\geqslant 0}$ 形成通常的马尔可夫链 (亦即 $r=1$ 阶的马尔可夫链).

14. (群上的随机游动) 设 G 为定义有二元运算 $\oplus$ 的有限群. 假设运算 $\oplus$ 满足通常的群性质:

(i) 如果 $x,y\in G$, 则 $x\oplus y\in G$.

(ii) 如果 $x,y,z\in G$, 则 $x\oplus(y\oplus z)=(x\oplus y)\oplus z$.

(iii) 存在单位元 e, 使得对一切 $x\in G$, 都有 $x\oplus e=e\oplus x=x$.

(iv) 对于每个 $x \in G$, 存在逆元素 $-x$, 使得 $x \oplus (-x) = (-x) \oplus x = e$.

设 $\xi_1, \xi_2, \cdots$ 是取值于 G 的随机元序列[①], 服从相同的概率 $Q(g) = \mathbf{P}\{\xi_n = g\}$, $g \in G$, $n \geqslant 0$.

令 $X_n = \xi_0 \oplus \xi_1 \oplus \cdots \oplus \xi_n$. 证明, 随机游动 $X = (X_n)_{n\geqslant 0}$ 形成马尔可夫链. 并给出转移概率矩阵.

15. (圆上的随机游动) 设 $\xi_1, \xi_2, \cdots$ 是独立同分布的取值于区间 $[0,1]$ 的随机变量序列, 具有连续的分布密度 $f(x)$. 令序列 $X = (X_n)_{n\geqslant 0}$, $X_0 = x \in [0,1)$ 和

$$X_n = x + \xi_1 + \cdots + \xi_n \qquad (\mathrm{mod}\ 1).$$

证明, 序列 $X = (X_n)_{n\geqslant 0}$ 形成以 $E = [0,1)$ 为状态空间的马尔可夫链. 并求其转移函数.

16. 设 $X = (X_n)_{n\geqslant 0}$ 与 $Y = (Y_n)_{n\geqslant 0}$ 是定义在同一个概率空间 $(\Omega, \mathscr{F}, \mathbf{P})$ 上的相互独立的两个马尔可夫链, 取值于可数集合 $E = \{i, j, \cdots\}$, 并具有相同的转移概率矩阵. 令 $X_0 = x$, $Y_0 = y$.

证明, 序列 $(X, Y) = (X_n, Y_n)_{n\geqslant 0}$ 形成马尔可夫链. 并求出其转移概率矩阵.

17. 设 $X_1, X_2, \cdots$ 是独立同分布的非负随机变量序列, 具有共同的连续分布函数. 引入如下的纪录时序列:

$$\mathscr{R}_1 = 1, \quad \mathscr{R}_k = \inf\{n \geqslant \mathscr{R}_{k-1}: \ X_n \geqslant \max(X_1, \cdots, X_{n-1})\}, \quad k \geqslant 2.$$

证明, 序列 $\mathscr{R} = (\mathscr{R}_k)_{k\geqslant 1}$ 形成马尔可夫链. 并求出其转移概率矩阵.

18. 设 $X_1, X_2, \cdots$ 是取值于离散集合的独立同分布的非负随机变量序列, $\mathscr{R} = (\mathscr{R}_k)_{k\geqslant 1}$ 是其记录时序列 (参阅上题). 令 $V_k = X_{\mathscr{R}_k}$, 则 $V = (V_k)_{k\geqslant 1}$ 是其记录值序列. 证明, $V = (V_k)_{k\geqslant 1}$ 形成马尔可夫链. 并求出其转移概率矩阵.

19. (关于在时间的变换之下马尔可夫性的保持) 设 $X = (X_n)_{0\leqslant n\leqslant N}$ 是具有可数状态空间 E 的不可约马尔可夫链, 转移概率矩阵为 $P = \| p_{ij} \|$ 和不变初始分布 $q = (q_i)$, 其中对任何 $i \in E$, 都有 $q_i > 0$ (关于不变分布可参阅《概率》第二卷 p257).

令 $\widetilde{X}_n = X_{N-n}$, 则序列 $\widetilde{X}^{(N)} = (\widetilde{X}_n)_{0\leqslant n\leqslant N}$ 是按时间 "反向的". 令 $\widetilde{P} = \| \widetilde{p}_{ij} \|$, 其中 $\widetilde{p}_{ij} = p_{ji}$. 证明, $\widetilde{P}$ 是随机的, 而且关于该矩阵, "反向" 序列是具有马尔可夫相依性 (参阅定义 2) 的随机变量序列.

注: 马尔可夫性的一种定义 (参阅定理 1 之 (7)) 说: 在给定 "现在" 的条件下, "将来" 与 "过去" 是相互独立的. 这种说法可以导致对按时间 "反向的" 序列 $X = (X_n)_{0\leqslant n\leqslant N}$ 的性状的一种看法. 由本题中的结果可以推出, 如果以不变

①此处似乎应当为相互独立的随机元序列 —— 译者注.

分布作为初始分布, 则 "反向的" 链仍然是马尔可夫链, 但一般来说, 转移概率矩阵是另外一个.

20. (可逆马尔可夫链) 设 $X = (X_n)_{n\geqslant 0}$ 是具有可数状态集合 E 的马尔可夫链, 具有移概率矩阵 $P = \|p_{ij}\|$ 和不变分布 $q = (q_i)$. 称这样的 (q, P) 马尔可夫链是可逆的 (例如, 可参阅, [124]), 如果对任何 $N \geqslant 1$, 按时间为 "反向的" 序列 $\widetilde{X}^{(N)} = (\widetilde{X}_n)_{0\leqslant n\leqslant N}$, 其中 $\widetilde{X}_n = X_{N-n}$, 仍然是 (q, P) 马尔可夫链.

证明, 不可约马尔可夫链是可逆的, 当且仅当, 对任何 $i, j \in E$, 都有:

$$q_i p_{ij} = q_j p_{ji}.$$

证明, 如果分布 $\lambda = (\lambda_i)$ $(\lambda_i \geqslant 0, \ \sum \lambda_i = 1)$ 和矩阵 P 满足平衡条件: 对任何 $i, j \in E$, 都有:

$$\lambda_i p_{ij} = \lambda_j p_{ji},$$

则 $\lambda = (\lambda_i)$ 重合于不变分布 $q = (q_i)$.

21. 证明, 在具有平稳分布 $q_i = C_N^i (1/2)^N$, $i = 0, 1, \cdots, N$ 的埃伦弗斯特模型 (第 8 节第 3 小节) 中, 有上述的平衡条件成立:

$$q_i p_{i,\, i+1} = q_{i+1} p_{i+1,\, i}.$$

(注意, 如果 $|i - j| > 1$, 考察模型 $P_{ij} = 0$.)

22. 设马尔可夫链 $X = (X_n)_{n\geqslant 0}$ 的转移概率矩阵为

$$P = \begin{pmatrix} 0 & 2/3 & 1/3 \\ 1/3 & 0 & 2/3 \\ 2/3 & 1/3 & 0 \end{pmatrix}.$$

证明, 它的不变分布为 $q = (1/3,\ 1/3,\ 1/3)$. 证明, 对于每个 $N \geqslant 1$, 序列 $\widetilde{X}^{(N)} = (X_{N-n})_{0\leqslant n\leqslant N}$ 都 (关于 P 的转置矩阵 $\widetilde{P}$) 是马尔可夫的. 证明, $X = (X_n)_{n\geqslant 0}$ 却不是可逆的. 试给出对于这一结果的直观解释.

23. 设 $X = (X_n)_{n\geqslant 0}$ 是 (严格意义下) 平稳的非退化高斯序列. 证明, 该序列具有马尔可夫性, 当且仅当, 对于一切 $n \geqslant 0$, $m \geqslant 0$, 协方差 $\mathrm{cov}(X_n, X_{n+m})$ 都具有如下形式:

$$\mathrm{cov}(X_n, X_{n+m}) = \sigma^2 \rho^m,$$

其中 $\sigma^2 > 0$, $|\rho| \leqslant 1$.

§2. 推广马尔可夫性和强马尔可夫性

1. 证明, 第 1 小节注中的函数 $\psi(x) = \mathbf{E}_x H$ 是 $\mathscr{E}$-可测的.

2. 证明性质 (12).

3. 证明性质 (13).

4. 第 3 小节例中的随机变量 $X_n - X_{\tau\wedge n}$ 与 $X_{\tau\wedge n}$ 之间是否具有独立性?

5. 证明 (23) 式.

6. 假设 (状态) 空间 E 为至多可数的, $(\Omega, \mathscr{F}) = (E^\infty, \mathscr{E}^\infty)$. 令 $\theta_n: \Omega \to \Omega$ 为移步算子, $n \geqslant 1$: 对于 $\omega = (x_0, x_1, \cdots)$, 有

$$\theta_n(\omega) = (x_n, x_{n+1}, \cdots).$$

用 $X = (X_n(\omega))_{n\geqslant 0}$ 表示标准的投影过程, 即对所有 $n \geqslant 0$ 与 $\omega = (x_0, x_1, \cdots)$, 有 $X_n(\omega) = x_n$.

对于一切 $\mathscr{F}$-可测的函数 $H = H(\omega)$, 我们令 (参阅 (1))

$$(H \circ \theta_n)(\omega) = H(\theta_n(\omega)), \quad n \geqslant 1,$$

而对于所有 $\mathscr{F}$-可测的集合 B, 则令 (试比较第五章第 1 节定义 2)

$$\theta_n^{-1}(B) = \{\omega: \theta_n(\omega) \in B\}, \quad n \geqslant 1.$$

试利用上述定义, 证明如下各性质:

(a) 对于 $m \geqslant 0$, $n \geqslant 1$, 有

$$X_m \circ \theta_n = X_{m+n},$$

亦即 $(X_m \circ \theta_n)(\omega) = X_{m+n}(\omega)$.

(b) 对于 $m \geqslant 0$, $n \geqslant 1$, 有

$$\theta_n^{-1}\{X_m \in A\} = \{X_m \circ \theta_n \in A\} = \{X_{m+n} \in A\},$$

亦即, 对于所有 $A \in \mathscr{E}$, 有

$$\theta_n^{-1}\{\omega: X_m(\omega) \in A\} = \{\omega: (X_m \circ \theta_n)(\omega) \in A\} = \{\omega: X_{m+n}(\omega) \in A\};$$

并且, 更加广义地有

$$\theta_n^{-1}\{X_0 \in A_0, \cdots, X_m \in A_m\} = \{X_0 \circ \theta_n \in A_0, \cdots, X_m \circ \theta_n \in A_m\}$$
$$= \{X_n \in A_0, \cdots, X_{m+n} \in A_m\}.$$

并且证明

$$\theta_n^{-1}(\mathscr{F}_m) = \sigma(X_n, \cdots, X_{m+n}) \tag{*}$$

(按照通常意义理解符号 $\theta_n^{-1}(\mathscr{F}_m)$ 与 $\sigma(X_n, \cdots, X_{m+n})$, 试给出相应的阐述).

7. 设 $H = H(\omega)$ 为 $\mathscr{F}$-可测的函数, $A \in \mathscr{B}(\mathbb{R})$. 证明

$$(H \circ \theta_n)^{-1}(A) = \theta_n^{-1}(H^{-1}(A)). \tag{**}$$

8. 设 $\tau = \tau(\omega)$ 为关于 $(\mathscr{F}_k)_{k\geqslant 0}$ 的停时, 其中 $\mathscr{F}_k = \sigma(X_0, X_1, \cdots, X_k)$ (即有限马尔可夫时). 试利用第 7 题与第 6 题中的性质 (**) 与 (*), 证明, 对于一切 $n \geqslant 0$, 时刻 $n + \tau \circ \theta_n$ 仍为停时, 亦即对一切 $m \geqslant n$, 都有 $\{\omega: n + (\tau \circ \theta_n)(\omega) = m\} \in \mathscr{F}_m$.

9. 设 $\sigma = \sigma(\omega)$ 为停时, $H = H(\omega)$ 为 $\mathscr{F}$-可测的函数. 我们将 $(H \circ \theta_\sigma)(\omega)$ 理解为函数 $H(\theta_{\sigma(\omega)}(\omega))$, 即当 $\omega \in \{\omega: \sigma(\omega) = n\}$ 时, 理解为变量 $H(\theta_n(\omega))$.

作为第 8 题的推广, 证明, $\sigma + \tau \circ \theta_\sigma$ 亦为停时.

10. 与上题中关于 $H \circ \theta_\sigma$ 的定义相对应, 对于停时 τ 与 σ, 我们将 $X_\tau \circ \theta_\sigma$ 理解为变量 $X_{\tau(\theta_\sigma(\omega))}(\theta_\sigma(\omega))$, 即当 $\omega \in \{\omega: \sigma(\omega) = n\}$ 时, 理解为变量 $X_{\tau(\theta_n(\omega))}(\theta_n(\omega))$.

作为第 6 题中性质 $X_m \circ \theta_n = X_{m+n}$ 的推广, 证明,

$$X_\tau \circ \theta_\sigma = X_{\tau\circ\theta_\sigma + \sigma}.$$

11. 设集合 $B \in \mathscr{E}$, 而

$$\tau_B(\omega) = \inf\{n \geqslant 0: X_n(\omega) \in B\}, \quad \sigma_B(\omega) = \inf\{n > 0: X_n(\omega) \in B\}$$

为序列 $X = (X_n)_{n\geqslant 0}$ 首次或零后首次属于集合 B 的时刻.

我们认为, 这两个时刻对于一切 $\omega \in \Omega$ 都是有限的. 再设 $\gamma = \gamma(\omega)$ 为停时. 证明, τ_B 与 σ_B 都是停时, 并且

$$\gamma + \tau_B \circ \theta_\gamma = \inf\{n \geqslant \gamma: X_n(\omega) \in B\}, \quad \gamma + \sigma_B \circ \theta_\gamma = \inf\{n > \gamma: X_n(\omega) \in B\}.$$

(这些关系式在变量 γ, τ_B, σ_B 取无限值时, 仍然正确, 此时集合 $\{\cdot\}$ 为空集. 试给出一般情况下的相应定义和所涉及的关系式.)

12. 设 τ 与 σ 是马尔可夫时. 证明, $\nu = \tau \circ \theta_\sigma + \sigma$ 也是马尔可夫时 (在集合 $\{\sigma = \infty\}$ 上, 约定 $\nu = \infty$).

13. 证明, 定理 2 中的严马尔可夫性 (7) 对于任何具有如下形式的马尔可夫时 $\tau \leqslant \infty$ 仍然成立:

$$\mathbf{E}_\pi[I(\tau < \infty)(H \circ \theta_\tau) \,|\, \mathscr{F}_\tau^X] = I(\tau < \infty)\mathbf{E}_{X_\tau} H \quad (\mathbf{P}_\pi\text{-a.s.})$$

(H 为有界或非负的 $\mathscr{F}$-可测的函数; 将 $\mathbf{E}_{X_\tau} H$ 理解为随机变量 $\Psi(X_\tau)$, 其中 $\Psi = \mathbf{E}_x H$).

证明, 如果 $K = K(\omega)$ 为 $\mathscr{F}_\tau$ 可测的函数, H 与 K 均有界或均非负, 则对任何马尔可夫时 $\tau \leqslant \infty$, 都有

$$\mathbf{E}_\pi[(I(\tau < \infty)K)(H \circ \theta_\tau)] = \mathbf{E}_\pi[(I(\tau < \infty)K)\mathbf{E}_{X_\tau} H].$$

14. 证明, 第 3 小节的例题中所考察的序列 $(X_{\tau\wedge n}, \mathbf{P}_x)_{n\geqslant 0}$, $x \in E$ 是马尔可夫链. 这一性质是否对于任何 (具有可数状态空间的) 马尔可夫链以及具有形式 $\tau = \inf\{n \geqslant 0:\ X_n \in A\}$ 的马尔可夫时都成立, 其中 $A \subseteq E$ (试比较 (15))?

15. 设函数 $H(x) = (Uh)(x)$ 是非负函数 $h = h(x)$ 的位势 (参阅附录第 7 节 p337). 证明, $H(x)$ 是方程 $V(x) = h(x) + TV(x)$(在非负函数类 $V = V(x)$ 中的)最小解.

16. 设 $y^\circ \in E$. 证明, 格林函数 $G(x, y^\circ)$ 是如下方程组的最小非负解:

$$V(x) = \begin{cases} 1 + TV(x), & x = y^\circ, \\ TV(x), & x \neq y^\circ. \end{cases}$$

17. 证明, 若 τ 与 σ 是两个马尔可夫时, 则对转移算子 T_n, $n \geqslant 0$, 有如下关系式成立:

$$T_\sigma T_\tau = T_{\sigma + \tau\circ\theta_\sigma}.$$

提示: 利用严马尔可夫性和第 10 题中的性质 $X_\tau \circ \theta_\sigma = X_{\tau\circ\theta_\sigma + \sigma}$.

18. 设集合 $D \in \mathscr{E}$, 而

$$\tau(D) = \inf\{n \geqslant 0:\ X_n \in D\}, \quad \sigma(D) = \inf\{n > 0:\ X_n \in D\}.$$

证明,

$$X_{\sigma(D)} = X_{\tau(D)\circ\theta_1} \quad \text{在}\{\sigma(D) < \infty\}\text{上},$$
$$T_{\sigma(D)} = TT_{\tau(D)}.$$

19. 设 $g \geqslant 0$, 而 $V_D(x) = T_{\tau(D)}g(x)$. 证明, $V_D(x)$ 是如下方程组的最小非负解:

$$V(x) = \begin{cases} g(x), & x \in D, \\ TV(x), & x \notin D. \end{cases}$$

特别地, 当 $g \equiv 1$ 时, 函数 $V_D(x) = \mathbf{P}_x\{\tau(D) < \infty\}$ 是如下方程组的最小非负解:

$$V(x) = \begin{cases} 1, & x \in D, \\ TV(x), & x \notin D. \end{cases}$$

20. 利用严马尔可夫性, 证明, 函数 $m_D(x) = \mathbf{E}_x\tau(D)$ 满足如下的方程组:

$$V(x) = \begin{cases} 0, & x \in D, \\ 1 + TV(x), & x \notin D. \end{cases}$$

再证明, $m_D(x)$ 是该方程组的最小非负解.

21. 设 $f = f(x)$ 是非负峰度函数. 证明, 该函数具有里斯分解:

$$f(x) = \widetilde{f}(x) + U\widetilde{h}(x),$$

其中 $\widetilde{f}(x)$ 是调和函数

$$\widetilde{f}(x) = \lim_n (T_n f)(x),$$

而 $U\widetilde{h}(x)$ 则是位势函数

$$\widetilde{h}(x) = f(x) - Tf(x).$$

22. 设 $X = (X_n)_{n\geqslant 1}$ 与 $Y = (Y_n)_{n\geqslant 1}$ 是两个相互独立的马尔可夫链, 具有相同的状态空间 $E = \{1, 2\}$ 与相同的转移概率矩阵 $\begin{pmatrix} \alpha & 1-\alpha \\ 1-\beta & \beta \end{pmatrix}$, 其中 $\alpha, \beta \in (0, 1)$. 令 $\tau = \inf\{n \geqslant 0:\ X_n = Y_n\}$(若 $\{\cdot\} = \varnothing$, 则 $\tau = \infty$) 为 "首次相遇" 时刻. 试求 τ 的概率分布.

23. 设 $X = (X_n, \mathscr{F}_n)_{n\geqslant 0}$ 是某个随机序列, $B \in \mathscr{B}(\mathbb{R})$. 根据例 6, $\tau_B = \inf\{n \geqslant 0:\ X_n \in B\}$ 与 $\sigma_B = \inf\{n > 0:\ X_n \in B\}$ 都是马尔可夫时. (始终认为 $\inf\varnothing = \infty$.) 证明, 在例 6 中有

$$\gamma_B = \sup\{0 \leqslant n \leqslant N:\ X_n \in B\}$$

(约定 $\sup\varnothing = 0$), 其中, N 是某个整数, 而非马尔可夫时.

24. 证明, 如果将函数 $H = H(\omega)$ 的有界性条件换为非负性, 则定理 1(以及定理 2) 中的断言仍然成立.

25. 设 $X = (X_n, \mathscr{F}_n)_{n\geqslant 0}$ 是马尔可夫序列, τ 为有限马尔可夫时. 令

$$\overline{X}_n = X_{n+\tau}, \quad \overline{\mathscr{F}}_n = \mathscr{F}_{n+\tau}.$$

证明, 随机序列 $\overline{X} = (\overline{X}_n, \overline{\mathscr{F}}_n)_{n\geqslant 0}$ 仍为马尔可夫序列, 并且具有与 $X = (X_n, \mathscr{F}_n)_{n\geqslant 0}$ 相同的转移函数. (这一性质可以视为严马尔可夫性的最简单形式.)

26. 设 $(X_1, X_2, \cdots)$ 是独立同分布的随机变量序列, 具有分布函数 $F = F(x)$. 令 $\mathscr{F}_0 = \{\varnothing, \Omega\}$, $\mathscr{F}_n = \sigma(X_1, \cdots, X_n)$, $n \geqslant 1$. 设 τ 是 (关于 $(\mathscr{F}_n)_{n\geqslant 0}$ 的) 马尔可夫时, 而 A 是 σ-代数 $\mathscr{F}_\tau$ 中的事件.

如果 τ 是有界的随机变量 $(0 \leqslant \tau(\omega) \leqslant T < \infty,\ \omega \in \Omega)$, 证明,

(a) 随机变量 $I_A, X_{1+\tau}, X_{2+\tau}, \cdots$ 相互独立.

(b) 每个随机变量 $X_{n+\tau}$ 都服从分布函数 $F = F(x)$, 即有 $\mathrm{Law}(X_{n+\tau}) = \mathrm{Law}(X_1)$, $n \geqslant 1$.

(由性质 (a) 与 (b) 可以推出: 随机变量序列 $(X_{1+\tau}, X_{2+\tau}, \cdots)$ 的概率结构与序列 $(X_1, X_2, \cdots)$ 的概率结构相同, 即 $\mathrm{Law}(X_{1+\tau}, X_{2+\tau}, \cdots) = \mathrm{Law}(X_1,$

$X_2,\cdots)$. 这一性质可以视为随机变量序列 $X=(X_n)_{n\geqslant 1}$ 的概率性质在 "时间的随机变换 $n\rightsquigarrow n+\tau$" 之下得以保持.)

证明, 对于任意的马尔可夫时 $\tau(0\leqslant\tau\leqslant\infty)$, 性质 (a) 转变为如下形式:

$$\mathbf{P}(A\cap\{\tau<\infty\};\ X_{1+\tau}\leqslant x_1,\cdots,X_{n+\tau}\leqslant x_n)=\mathbf{P}(A\cap\{\tau<\infty\})F_1(x)\cdots F_n(x)$$

(对一切 $n\geqslant 1$ 与 $x_n\in\mathbb{R}$ 成立).

提示: 注意对于问题的解利用关系式

$$\begin{aligned}\mathbf{P}(A\cap\{\tau<\infty\};\ &X_{1+\tau}\leqslant x_1,\cdots,X_{n+\tau}\leqslant x_n)\\&=\sum_{k=0}^{\infty}\mathbf{P}(A\cap\{\tau=k\};\ X_{1+k}\leqslant x_1,\cdots,X_{n+k}\leqslant x_n)\end{aligned}$$

其中, 事件 $A\cap\{\tau=k\}$ 与 $\{X_{1+k}\leqslant x_1,\cdots,X_{n+k}\leqslant x_n\}$ 是相互独立的.

§3. 马尔可夫链的极限、遍历和平稳概率分布问题

1. 试举例说明, 存在这样的马尔可夫链, 其极限 $\pi_j=\lim\limits_n p_{ij}^{(n)}$

(a) 与初始状态 j 无关.

(b) 依赖于初始状态 j.

2. 分别举出具有遍历性的链与不具有遍历性的链的例子.

3. 试举出平稳分布不是遍历的例子

4. 试举出这样的转移概率矩阵的例子, 对其任何状态空间上的概率分布都是平稳分布.

§4. 马尔可夫链的状态按转移概率矩阵的代数性质分类

1. 试对第 1 小节末尾所描述的关于非本质状态或本质状态的定义, 补充给出以转移概率 $p_{ij}^{(n)}$, $i,j\in E$, $n\geqslant 1$ 的性质作为术语的数量公式.

2. 设 $\mathbb{P}$ 是某个具有有限状态的不可约马尔可夫链的转移概率矩阵, 有 $\mathbb{P}^2=\mathbb{P}$. 试研究该矩阵 $\mathbb{P}$ 的结构.

3. 设 $\mathbb{P}$ 是有限马尔可夫链 $X=(X_n)_{n\geqslant 0}$ 的转移概率矩阵. 设 $\sigma_1,\sigma_2,\cdots$ 是与 X 独立的非负整数值的独立同分布随机变量序列. 令 $\tau_0=0$, $\tau_n=\sigma_1+\cdots+\sigma_n$, $n\geqslant 1$, 并令 $\widetilde{X}_n=X_{\tau_n}$. 证明, 序列 $\widetilde{X}=(\widetilde{X}_n)_{n\geqslant 0}$ 是马尔可夫链. 试求其转移概率矩阵 $\widetilde{\mathbb{P}}$. 并且证明, 如果状态 i,j 对于 X 是连通的, 那么它们对于 $\widetilde{X}$ 也是连通的.

4. 设 $X=(X_n)_{n\geqslant 0}$ 是以 $E=\{0,1\}$ 为状态空间, 以 $\mathbb{P}=\begin{pmatrix}\alpha & 1-\alpha\\ 1-\beta & \beta\end{pmatrix}$ 为转移概率矩阵的马尔可夫链, 其中 $\alpha,\beta\in(0,1)$. 令马尔可夫时

$$\nu=\inf\{n\geqslant 1:\ X_{n-1}=X_n=0\}.$$

证明,

$$\mathbf{E}_0\nu=\frac{2-(\alpha+\beta)}{\alpha(1-\beta)}.$$

5. 我们来考察具有 3 个状态 $E=\{1,2,3\}$ 和转移概率矩阵

$$\mathbb{P}=\begin{pmatrix}\alpha & 1-\alpha & 0\\ 0 & \beta & 1-\beta\\ 1-\gamma & 0 & \gamma\end{pmatrix}$$

的马尔可夫链, 其中 $\alpha,\beta,\gamma\in(0,1)$. 证明, 具有该种转移概率矩阵的马尔可夫链是不可约的. 试讨论此类马尔可夫链的平稳分布问题.

6. 试问, 在如下两种情况下, 马尔可夫链的所有状态可否都是非本质的?

(a) 共有有限个状态.

(b) 共有可数个状态.

§5. 马尔可夫链的状态按转移概率矩阵的渐近性质分类

1. 考察以 $0,1,2,\cdots$ 为状态的不可约的链. 证明, 为了使其为非自返的, 必须且只需, 方程组 $u_j=\sum\limits_i u_i p_{ij}$, $j=0,1,\cdots$ 有满足条件 $u_i\neq\text{const}$ 的有界解, 其中 $i=0,1,\cdots$.

2. 证明, 为了使得以 $0,1,\cdots$ 为状态的不可约的链是自返的, 只需存在序列 $(u_0,u_1,\cdots)$, 有 $u_i\to\infty$, $i\to\infty$, 使得对一切 $j\neq 0$, 都有 $u_j\geqslant\sum\limits_i u_i p_{ij}$.

3. 证明, 为了使得以 $0,1,\cdots$ 为状态的不可约的链是自返的和正的, 必须且只需, 方程组 $u_j=\sum\limits_i u_i p_{ij}$, $j=0,1,\cdots$, 有不恒等于 0 的解, 其中 $\sum\limits_i |u_i|<\infty$.

4. 设马尔可夫链具有状态 $0,1,\cdots$ 与转移概率:

$$p_{00}=r_0,\quad p_{01}=p_0>0,$$

$$p_{ij}=\begin{cases}p_i>0, & j=i+1,\\ r_i\geqslant 0, & j=i,\\ q_i>0, & j=i-1,\\ 0, & \text{其余情形}.\end{cases}$$

令 $\rho_0=1$, $\rho_m=\frac{q_1\cdots q_m}{p_1\cdots p_m}$. 证明如下各断言:

$$\text{链为自返的} \quad \Leftrightarrow \quad \sum \rho_m=\infty,$$
$$\text{链为非自返的} \quad \Leftrightarrow \quad \sum \rho_m<\infty,$$
$$\text{链为自返与正的} \quad \Leftrightarrow \quad \sum \rho_m=\infty,\quad \sum \frac{1}{p_m\rho_m}<\infty,$$
$$\text{链为自返与零的} \quad \Leftrightarrow \quad \sum \rho_m=\infty,\quad \sum \frac{1}{p_m\rho_m}=\infty.$$

5. 证明,

$$f_{ik}\geqslant f_{ij}f_{jk},$$
$$\sup_n p_{ij}^{(n)}\leqslant f_{ij}\leqslant \sum_{n=1}^{\infty}p_{ij}^{(n)}.$$

6. 证明, 对于任何具有可数状态集合的马尔可夫链, $p_{ij}^{(n)}$ 都存在切萨罗(Cesàro) 意义下的极限:

$$\lim_n \frac{1}{n}\sum_{k=1}^{n}p_{ij}^{(k)}=\frac{f_{ij}}{\mu_j}.$$

7. 设 $\eta_1,\eta_2,\cdots$ 为独立同分布的随机变量序列, 有 $\mathbf{P}\{\eta_k=j\}=p_j$, $j=0,1,\cdots$. 我们来考察马尔可夫链 $\xi_0,\xi_1,\cdots$, 其中 $\xi_{k+1}=(\xi_k)^+\eta_{k+1}$, $k\geqslant 0$. 试写出它的转移概率矩阵, 并证明, 如果 $p_0>0$, $p_0+p_1<1$, 则链为自返的, 当且仅当, $\sum\limits_k kp_k\leqslant 1$.

8. 令 $\sigma_i=\inf\{n>0: X_n=i\}$ ($\sigma_i=\infty$, 如果 $\{\cdot\}=\varnothing$), $\sigma_i^0=\sigma_i$, 且对 $n\geqslant 1$, 令

$$\sigma_i^n=\begin{cases}\sigma_i^{n-1}+\sigma_i\circ\theta_{\sigma_i^{n-1}}, & \text{如果 } \sigma_i^{n-1}<\infty,\\ \infty, & \text{如果 } \sigma_i^{n-1}=\infty.\end{cases}$$

证明,

$$\mathbf{P}_i\{\sigma_i^n<\infty\}=(\mathbf{P}_i\{\sigma_i<\infty\})^n \quad (=f_{ii}^n).$$

9. 以 $N_{\{i\}}$ 表示马尔可夫链访问状态 i 的次数.

(a) 证明,

$$\mathbf{E}_i N_{\{i\}}=\frac{1}{1-\mathbf{P}_i\{\sigma_i<\infty\}} \quad \left(=\frac{1}{1-f_{ii}}\right).$$

(b) 运用术语 $\mathbf{E}_i N_{\{i\}}$ 改写关于状态 $i\in E$ 的自返性与非自返性的判别条件.

(c) 证明,

$$\mathbf{E}_i N_{\{i\}}=\mathbf{P}_i\{\sigma_i<\infty\}\cdot\mathbf{E}_i N_{\{i\}}.$$

10. (非自返的充分必要条件) 设 $X=(X_n)_{n\geqslant 0}$ 是具有可数状态集合 E 与转移概率矩阵 $\|p_{xy}\|$ 的不可约马尔可夫链. 证明, 使得该链非自返的充分必要条件是: 存在 (非平凡的) 有界函数 $f=f(x)$ 与状态 $x^\circ\in E$, 使得

$$f(x)=\sum_{y\neq x^\circ}p_{xy}f(y),\quad x\neq x^\circ$$

(即在集合 $E\backslash\{x^\circ\}$ 上的调和性).

11. (非自返的充分条件) 设 $X=(X_n)_{n\geqslant 0}$ 是具有可数状态集合 E 的不可约马尔可夫链. 假设存在有界函数 $f=f(x)$ 使得对某个 $B\subseteq E$, 有

$$f(x^\circ)<f(x) \quad \text{对某个 } x^\circ\in B \text{与一切 } x\in B$$

和

$$\sum_{y\in E}p_{xy}f(y)\leqslant f(x),\quad x\in\overline{B}\ \ (=E\backslash B)$$

(在集合 $\overline{B}$ 上的上调和性).

证明, 该条件对于链的非自返性是充分的.

12. (自返的充分条件) 设 $X=(X_n)_{n\geqslant 0}$ 是具有可数状态集合 E 的不可约马尔可夫链. 假设存在函数 $h=h(x)$, $x\in E$, 使得对一切常数 c, 集合 $B_c=\{x: h(x)<c\}$ 都是有限的, 并且对于某个有限集合 $A\subseteq E$, 有

$$\sum_{y\in E}p_{xy}h(y)\leqslant h(x),\quad x\in\overline{A}$$

(在集合 $\overline{A}$ $(=E\backslash A)$ 上的上调和性).

证明, 该链是自返的.

13. 证明, 上题中的条件也是必要的.

14. 设 $(\xi_n)_{n\geqslant 1}$ 是取值于 $\mathbb{R}$ 的独立同分布随机变量序列, $X=(X_n)_{n\geqslant 0}$ 为序列 $(\xi_n)_{n\geqslant 1}$ 所构成的随机游动: $X_0=0$, $X_n=\xi_1+\cdots+\xi_n$. 以 N_B 表示随机游动 X 访问博雷尔集合 B 的次数 (亦即 $N_B=\sum\limits_{n\geqslant 0}I_B(X_n)$), 而 $U(B)=\mathbf{E}\,N_B$ 是 N_B 的数学期望, 在附录中 (第 7 节 p337) 将这一测度称为 (关于初始状态 $x=0$ 的)位势测度.

类似于对具有可数状态集合马尔可夫链的自返性与非自返性的定义 (参阅第 1 小节中的定义 1 和定义 2), 如果对于任何有限区间 $I\subseteq\mathbb{R}$, 都有

$$U(I)=\infty,$$

我们说随机游动 (它当然是取值于 $\mathbb{R}$ 的马尔可夫链) 是自返的; 如果对于任何有限区间 $I\subseteq\mathbb{R}$, 都有

$$U(I)<\infty,$$

则随机游动是非自返的.

以下假设数学期望 $\mathbf{E}\xi_1$ 确定.

证明, 必有如下三种情况之一成立:

1) $X_n \to \infty$ ($\mathbf{P}$-a.s.) 且随机游动非自返.

2) $X_n \to -\infty$ ($\mathbf{P}$-a.s.) 且随机游动非自返.

3) $\underline{\lim} X_n = -\infty$, $\overline{\lim} X_n = +\infty$ (即在 $-\infty$ 与 $+\infty$ 之间随机振荡), 此时既可能自返; 也可能非自返.

15. 在第 14 题中的条件下 (记 $\mu = \mathbf{E}\xi_1$), 证明,

1) 若 $0 < \mu \leqslant \infty$, 则 $X_n \to \infty$ ($\mathbf{P}$-a.s.).

2) 若 $-\infty \leqslant \mu < 0$, 则 $X_n \to -\infty$ ($\mathbf{P}$-a.s.).

3) 若 $\mu = 0$, 则 $\underline{\lim} X_n = -\infty$, $\overline{\lim} X_n = +\infty$, 且随机游动自返.

16. 证明, 随机游动 $X = (X_n)_{n\geqslant 0}$ 是非自返的, 当且仅当, 概率为 1 地有 $|X_n| \to \infty$, $n \to \infty$.

17. 我们来考察这样的马尔可夫链 $X = (X_n)_{n\geqslant 0}$, 其转移概率 p_{ij}, $i, j \in E = \{0, \pm1, \pm2, \cdots\}$ 满足条件: 若 $|i-j| > 1$, 则 $p_{ij} = 0$, 且在等式

$$p_{i,i-1} + p_{ii} + p_{i,i+1} = 1, \quad i \in E$$

中, 三个加项均为正.

证明, 这样的链是不可约的和非周期的. 试找出使得这样的链为自返的、非自返的、正的自返的、零的关于转移概率 (p_{ii}, $p_{i,i-1}$, $p_{i,i+1}$, $i \in E$) 的条件 (试比较第 4 题).

提示: 研究概率 $V(i) = \mathbf{P}_i\{\tau_{j^\circ} = \infty\}$, $i \in E$, 所满足的结构关系式 (对于每个固定的 j°).

18. (关于高尔顿–沃森过程中所表述的概率) 为了研究家族在英国的消失规律, 高尔顿与沃森 (于 19 世纪 70 年代) 提出了如下的以他们的姓氏冠名的模型:

设 $\xi_0, \xi_1, \xi_2, \cdots$ 为取值于集合 $\mathbb{N} = \{0, 1, 2, \cdots\}$ 的随机变量序列, 它们按照 "随机个随机变量的和" 的原则定义:

$$\xi_{n+1} = \eta_1^{(n)} + \cdots + \eta_{\xi_n}^{(n)}, \quad n \geqslant 0.$$

(试比较第一章第 12 节例 4.) 其中, $\{\eta_i^{(n)}, i \geqslant 1, n \geqslant 0\}$ 是相互独立的随机变量族, 其中每个随机变量都与随机变量 η 同分布, 即有 $\mathbf{P}\{\eta = k\} = p_k$, $k \geqslant 0$, $\sum\limits_{k=1}^{\infty} p_k = 1$. 将 ξ_n 理解为第 n 代 "父亲" 的数目, 而 $\eta_i^{(n)}$ 为其中第 i 个 "父亲" 的 "男孩" 的数目. 于是, ξ_{n+1} 就是第 $n+1$ 代全体 "男孩" 的数目. (如果 $\xi_n = 0$, 则认为对一切 $k > n$, 都有 $\xi_k = 0$.)

设 $\tau = \inf\{n \geqslant 0 : \xi_n = 0\}$ 为该家族消失的时刻 (如同往常, 如果对一切 $n \geqslant 0$, 都有 $\xi_n > 0$, 则认为 $\tau = \infty$). 由此所产生出的第一个问题便是: 概率

$$q = \mathbf{P}\{\tau < \infty\},$$

即在有限时间内家族消失的概率是怎样的?

回答这个问题的最有力的方法是母函数方法 (参阅第二章第 6 节第 28 题与附录第 3 节). 令 $g(s) = \sum\limits_{k=0}^{\infty} p_k s^k$, $|s| \leqslant 1$, 是取值 $0, 1, 2, \cdots$ 的随机变量 η 的母函数 ($g(s) = \mathbf{E}s^\eta$), 而 $f_n(s) = \mathbf{E}s^{\xi_n}$ 是随机变量 ξ_n 的母函数.

证明, 对于高尔顿–沃森模型, 有:

(a) $f_n(s) = f_{n-1}(g(s)) = f_{n-2}(g(g(s))) = \cdots = f_0(g^{(n)}(s))$, 其中 $g^{(n)} = (g \circ \cdots \circ g)(s)$ (复合 n 次).

(b) 若 $\xi_0 = 1$, 则 $f_0(s) = s$, $f_n(s) = g^{(n)}(s) = g(f_{n-1}(s))$.

(c) $f_n(0) = \mathbf{P}\{\xi_n = 0\}$.

(d) 事件 $\{\xi_n = 0\} \subseteq \{\xi_{n+1} = 0\}$.

(e) $\mathbf{P}\{\tau < \infty\} = \mathbf{P}\left(\bigcup\limits_{n=1}^{\infty} \{\xi_n = 0\}\right) = \lim\limits_{N\to\infty} \mathbf{P}\{\xi_N = 0\} = \lim\limits_{N\to\infty} f_N(0)$.

(f) 若 $q = \mathbf{P}\{\tau < \infty\}$, 则 q 是方程 $x = g(x)$ 的根, 其中 $0 \leqslant x \leqslant 1$.

19. (续第 18 题) 设 $g(s) = \mathbf{E}s^\eta$ 是取值 $0, 1, 2, \cdots$ 的随机变量 η 的母函数. 证明,

(a) $g = g(s)$ 是区间 $[0, 1]$ 中的非降下凸函数.

(b) 如果 $\mathbf{P}\{\eta = 0\} < 1$, 则函数 $g = g(s)$ 是严格上升的.

(c) 如果 $\mathbf{P}\{\eta \leqslant 1\} < 1$, 则函数 $g = g(s)$ 是严格凸的.

(d) 如果 $\mathbf{P}\{\eta \leqslant 1\} < 1$ 且 $\mathbf{E}\eta \leqslant 1$, 则方程 $x = g(x)$, $0 \leqslant x \leqslant 1$, 有唯一的根 $q \in [0, 1]$.

(e) 如果 $\mathbf{P}\{\eta \leqslant 1\} < 1$ 且 $\mathbf{E}\eta > 1$, 则方程 $x = g(x)$, $0 \leqslant x \leqslant 1$, 有两个根: $x = 1$ 与 $q \in (0, 1)$.

提示: 首先证明, 对于 $x \in [0, 1]$, 有 $g'(x) \geqslant 0$ 与 $g''(x) \geqslant 0$. 为证明所陈述的断言, 可以分别在 $\mathbf{E}\eta \leqslant 1$ 与 $\mathbf{E}\eta > 1$ 的情形下考察有界函数 $g = g(x)$.

20. (续第 18 题与第 19 题) 设在高尔顿–沃森模型中有 $\mathbf{E}\eta > 1$. 证明, 此时家族消失的概率 $q = \mathbf{P}\{\tau < \infty\}$ 是方程 $x = g(x)$ 的唯一根, 该根严格地介于 0 与 1 之间, 亦即

$$\mathbf{E}\eta > 1 \quad \Rightarrow \quad 0 < \mathbf{P}\{\tau < \infty\} < 1.$$

而若 $\mathbf{E}\eta \leqslant 1$ 且 $p_1 \neq 1$, 则家族消失的概率等于 1, 即

$$\mathbf{E}\eta \leqslant 1 \quad \Rightarrow \quad \mathbf{P}\{\tau < \infty\} = 1.$$

21. 设在高尔顿–沃森模型中有 $p_1 < 1$. 证明, 对于一切 $k \geqslant 1$, 都有 $\mathbf{P}\{\xi_n = k \text{ i.o.}\} = 0$. 并由此断言 $\mathbf{P}\left\{\lim\limits_n \xi_n \in \{0, \infty\}\right\} = 1$.

22. 设 $(X_n)_{n\geqslant 0}$ 是具有可数状态集合 $E=\{i,j,\cdots\}$ 的马尔可夫链. 证明, 如果其所有状态都是非本质的, 则该链为不可约的与自返的, 当且仅当:

(a) 对一切 $i,j\in E$, 都有 $f_{ij}=1$(亦即 $\mathbf{P}_i\{\sigma(j)<\infty\}=1$, 其中 $\sigma(j)=\inf\{n>0:\ X_n=j\}$).

(b) 每个峰态的有限非负函数 $h=h(i)$, $i\in E$(即 $h(i)\geqslant\sum\limits_{j\in E}p_{ij}h(j)$, $i\in E$, 其中 $\|p_{ij}\|$ 是链的转移概率矩阵), 都是常数.

提示: (b) 的必要性证明见附录第 7 节 p347. 为证充分性, 只需对任何 $i,j\in E$, 证明

$$f_{ij}=p_{ij}+\sum_{k\neq j}p_{ik}f_{kj},$$

并由此导出, 如果峰态函数都是常数, 则对一切 $i,j\in E$, 都有 $f_{ij}=1$, 再由 (a), 知其等价于链的不可约性与自返性.

§6. 7. 可数与有限马尔可夫链的极限分布、遍历分布和平稳分布

1. 试考察具有下述转移概率矩阵的马尔可夫链的平稳、极限和遍历分布:

$$\mathbb{P}=\begin{pmatrix}1/2 & 0 & 1/2 & 0\\ 0 & 0 & 0 & 1\\ 1/4 & 1/2 & 1/4 & 0\\ 0 & 1/2 & 1/2 & 0\end{pmatrix}.$$

2. 设 $\mathbb{P}=\|p_{ij}\|$ 为有限二阶随机矩阵 (即对 $i=1,\cdots,m$, 有 $\sum\limits_{j=1}^{m}p_{ij}=1$; 对 $j=1,\cdots,m$, 有 $\sum\limits_{i=1}^{m}p_{ij}=1$). 对于相应的马尔可夫链, 其平稳分布是向量 $\mathbb{Q}=(1/m,\cdots,1/m)$.

3. 设 $X=(X_n)_{n\geqslant 0}$ 是具有两个状态的马尔可夫链, 有 $E=\{0,1\}$ 与转移概率矩阵 $\mathbb{P}=\begin{pmatrix}\alpha & 1-\alpha\\ 1-\beta & \beta\end{pmatrix}$, 其中 $0<\alpha<1$, $0<\beta<1$.

证明,

(a) $\mathbb{P}^n=\dfrac{1}{2-(\alpha+\beta)}\begin{pmatrix}1-\beta & 1-\alpha\\ 1-\beta & 1-\alpha\end{pmatrix}+\dfrac{(\alpha+\beta-1)^n}{2-(\alpha+\beta)}\begin{pmatrix}1-\alpha & -(1-\alpha)\\ -(1-\beta) & 1-\beta\end{pmatrix}$.

(b) 若初始分布为 $\pi=(\pi(0),\ \pi(1))$, 则

$$\mathbf{P}\{X_n=0\}=\frac{1-\beta}{[illegible]}+(\alpha+\beta-1)^n\left[\pi(0)-\frac{1-\beta}{[illegible]}\right]$$

4. (续第 3 题) 试求平稳分布 π°, 并计算

$$\mathbf{cov}_{\pi^\circ}(X_n,\ X_{n+1})=\mathbf{E}_{\pi^\circ}X_nX_{n+1}-\mathbf{E}_{\pi^\circ}X_n\,\mathbf{E}_{\pi^\circ}X_{n+1}.$$

证明, 若 $S_n=X_1+\cdots+X_n$, 则

$$\mathbf{E}_{\pi^\circ}S_n=\frac{n(1-\alpha)}{2-(\alpha+\beta)}\quad 且\quad \mathbf{D}_{\pi^\circ}S_n\leqslant cn,$$

其中 c 为某个常数.

再证明, (相对于测度 $\mathbf{P}_0$, $\mathbf{P}_1$ 与 $\mathbf{P}_{\pi^\circ}$) 几乎必然地有

$$\frac{S_n}{n}\to\frac{1-\alpha}{2-(\alpha+\beta)}.$$

5. 设 $\mathbb{P}=\|p_{ij}\|$ 是转移概率矩阵 $(i,j\in E=\{0,1,2,\cdots\})$, 对一切 $i\in E$, 都有 $p_{i,i+1}=p_i$, $p_{i0}=1-p_i$, 其中 $0<p_i<1$, $i\geqslant 1$, 而 $p_0=1$.

证明, 具有这样的转移概率矩阵的马尔可夫链的所有状态都是自返的, 当且仅当, $\lim\limits_n\prod\limits_{j=1}^{n}p_j=0$(或, 等价地有 $\sum\limits_{j=1}^{\infty}(1-p_j)=\infty$).

证明, 若所有状态都是自返的, 则它们都是正自返的, 当且仅当,

$$\sum_{k=1}^{\infty}\prod_{j=1}^{k}p_j<\infty.$$

6. 证明, 若 $X=(X_k)_{k\geqslant 0}$ 是不可约正自返的马尔可夫链, 具有不变分布 π°, 则对所有固定 $x\in E$, (对于任何初始分布 π, 都 $\mathbf{P}_\pi$-a.s. 地有)

$$\frac{1}{n}\sum_{k=0}^{n-1}I_{\{x\}}(X_k)\to\pi^\circ(x),\quad n\to\infty,$$

与

$$\frac{1}{n}\sum_{k=0}^{n-1}p_{yx}^{(k)}\to\pi^\circ(x),\quad n\to\infty,\quad 任何\ y\in E.$$

(试比较第一章第 12 节中的大数定律.)

再证明, 如果马氏链是不可约的、自返的和零的, 则对任何固定的 $x\in E$, (对于任何初始分布 π, 都 $\mathbf{P}_\pi$-a.s. 地有)

$$\frac{1}{n}\sum_{k=0}^{n-1}I_{\{x\}}(X_k)\to 0,\quad n\to\infty,$$

与

$$\frac{1}{n}\sum_{k=0}^{n-1}p_{yx}^{(k)}\to 0,\quad n\to\infty,\quad 任何\ y\in E.$$

7. 设 $(X_n)_{n\geqslant 0}$ 是状态空间为有限集合 $E=\{0,1,\cdots,N\}$ 的马尔可夫链, 且该链又是鞅. 证明,

(a) 状态 $\{0\}$ 与 $\{N\}$ 为吸引态 $(p_{00}=p_{NN}=1)$.

(b) 若 $\tau(x)=\inf\{n\geqslant 0:\ X_n=x\}$, 则 $\mathbf{P}_x\{\tau(N)<\tau(0)\}=x/N$.

§8. 作为马尔可夫链的简单随机游动

1. 利用如下所述的概率方法证明斯特林公式 ($n!\sim\sqrt{2\pi}n^{n+1/2}e^{-n}$, 见 [10] 第 27.18 题): 令 $S_n=X_1+\cdots+X_n$, $n\geqslant 1$, 其中 $X_1,X_2,\cdots$ 是独立的随机变量序列, 服从参数为 $\lambda=1$ 的泊松分布. 依次证明,

$$\text{(a)}\quad \mathbf{E}\left(\frac{S_n-n}{\sqrt{n}}\right)^-=e^{-n}\sum_{k=0}^{n}\left(\frac{n-k}{\sqrt{n}}\right)\frac{n^k}{k!}=\frac{n^{n+1/2}e^{-n}}{n!}.$$

$$\text{(b)}\quad \operatorname{Law}\left[\left(\frac{S_n-n}{\sqrt{n}}\right)^-\right]\to\operatorname{Law}[N^-],$$

其中 N 是正态分布的随机变量.

$$\text{(c)}\quad \mathbf{E}\left[\left(\frac{S_n-n}{\sqrt{n}}\right)^-\right]\to\mathbf{E}[N^-]=\frac{1}{\sqrt{2\pi}}.$$

$$\text{(d)}\quad n!\sim\sqrt{2\pi}n^{n+1/2}e^{-n}.$$

2. 证明马尔可夫性 (28).

3. 证明 (30) 式.

4. 证明, 埃伦弗斯特模型中的马尔可夫链的所有状态都是自返的.

5. 验证 (31) 式与 (32) 式.

6. 证明, 对于 $\mathbb{Z}=\{0,\pm1,\pm2,\cdots\}$ 上的最简单的随机游动, 即 $p_{x,x+1}=p$, $p_{x,x-1}=1-p$, 函数 $f(x)=\left(\frac{1-p}{p}\right)^x$, $x\in\mathbb{Z}$, 是调和的.

7. 设 $\xi_1,\cdots,\xi_n$ 为独立同分布的随机变量, 令 $S_k=\xi_1+\cdots+\xi_k$, $k\leqslant n$. 证明,

$$\sum_{k\leqslant n}I(S_k>0)\overset{\mathrm{d}}{=}\min\{1\leqslant k\leqslant n:\ S_k=\max_{j\leqslant n}S_j\},$$

其中, $\overset{\mathrm{d}}{=}$ 表示同分布. (斯帕尔–安德森的这一结果补充说明了, 为什么在伯努利场合下, 位于正半轴的时间与达到最大值的时间都服从相同的规律 —— 反正弦律; 试比较第一章第 10 节以及该节中的第 4 题与第 5 题.)

8. 设 $\xi_1,\xi_2,\cdots$ 为独立的伯努利随机变量序列, 有 $\mathbf{P}\{\xi_n=1\}=\mathbf{P}\{\xi_n=-1\}=1/2$, $n\geqslant 1$. 令 $S_0=0$, $S_n=\xi_1+\cdots+\xi_n$. 证明, 若 τ 为马尔可夫停时, 且

$$\overline{S}_n=S_{n\wedge\tau}-(S_n-S_{n\wedge\tau})=\begin{cases}S_n, & n\geqslant\tau,\\ 2S_\tau-S_n, & n<\tau,\end{cases}$$

则有 $(\overline{S}_n)_{n\geqslant0}\overset{\mathrm{d}}{=}(S_n)_{n\geqslant0}$, 亦即序列 $(\overline{S}_n)_{n\geqslant0}$ 的概率分布与序列 $(S_n)_{n\geqslant0}$ 的概率分布重合. (本结果称为关于对称简单随机游动 $(S_n)_{n\geqslant0}$ 的*安德烈反射原理*, 试比较第一章第 10 节中关于反射原理的其他版本.)

9. 设 $X=(X_n)_{n\geqslant0}$ 是 d 维格子点集 $\mathbb{Z}^d$ 上的简单随机游动, 即有 $X_0=0$,

$$X_n=\xi_1+\cdots+\xi_n,\quad n\geqslant 1,$$

其中 $\xi_1,\xi_2,\cdots$ 是独立同分布的 d 维随机向量序列, 有

$$\mathbf{P}\{\xi_i=\mathbf{e}\}=\frac{1}{2d},$$

$e=(e_1,\cdots,e_d)$ 为 $\mathbb{R}^d$ 中的标准基向量, 即 $e_i=0,-1,1$, 且 $|\mathbf{e}|\equiv|e_1|+\cdots|e_d|=1$.

设 A 是 $\mathbb{R}^d$ 中的以 $\mathbf{0}=(0,\cdots,0)$ 为球心的开球. 证明多元中心极限定理的如下形式:

$$\lim_n\mathbf{P}\left\{\frac{X_n}{\sqrt{n}}\in A\right\}=\int_A\left(\frac{d}{2\pi}\right)^d e^{-\frac{d|x|^2}{2}}dx_1\cdots dx_d.$$

提示: 首先证明 ξ_1 的特征函数 $\varphi(\mathbf{t})=\mathbf{E}\,e^{i(\mathbf{t},\xi_1)}$, $\mathbf{t}=(t_1,\cdots,t_d)$ 为 $\varphi(\mathbf{t})=d^{-1}\sum\limits_{j=1}^{d}\cos t_j$. 然后利用多维场合下的连续性定理 (参阅第三章第 3 节定理 1); 亦可比较第三章第 3 节第 5 题.

10. 设 $X=(X_n)_{n\geqslant0}$ 是上面第 9 题中的随机游动, $N_n=\sum\limits_{k=0}^{n-1}I(X_k=0)$ 是使得 $X_k=0$ 的 $k\in\{0,1,\cdots,n-1\}$ 的个数. 第七章第 9 节证明了, 当 $d=1$ 时, 有 $\mathbf{E}\,N_n\sim\sqrt{\frac{2}{\pi}n}$, $n\to\infty$. (《概率》第二卷第七章第 9 节 (17) 与 (18) 式中的 $\frac{1}{2\pi}$ 均应为 $\frac{2}{\pi}$.)

(a) 证明, 对于 $d\geqslant 2$, 有如下的关系式成立:

$$\mathbf{E}\,N_n\sim\begin{cases}\dfrac{1}{\pi}\ln n, & d=2,\\ c_d, & d\geqslant 3,\end{cases}$$

其中常数 $c_d=1/\mathbf{P}\{\sigma_d=\infty\}$, 而 $\sigma_d=\inf\{k>0:\ X_k=0\}$(如果 $\{\cdot\}=\varnothing$, 则 $\sigma_d=\infty$). 试求常数 c_d.

(b) 证明, 当 $d=2$ 时, 有

$$\lim_n \mathbf{P}\left\{\frac{N_n}{\ln n} \geqslant x\right\} = e^{-\pi x}, \quad x>0,$$

与

$$\mathbf{P}\{\sigma_1 > n\} = \mathbf{P}\{N_n = 0\} \sim \frac{\pi}{\ln n}, \quad n \to \infty.$$

(c) 证明, 当 $d \geqslant 3$, $n \to \infty$ 时, 有

$$\mathbf{P}\{\sigma_1 = 2n\} \sim \frac{\mathbf{P}\{X_{2n}=0\}}{\left(1+\sum\limits_{k=1}^{\infty}\mathbf{P}\{X_{2k}=0\}\right)^2}.$$

(d) 证明, 对任何 $d \geqslant 1$, 都有

$$\mathbf{P}\{\sigma_1 = \infty\} = \frac{1}{1+\sum\limits_{k=1}^{\infty}\mathbf{P}\{X_{2k}=0\}}.$$

注: 关于性质 (d) 的证明, 已经本质上包含在第八章第 5 节定理 1 的证明之中. 指出如下事实是有益的: 由 (d) 中的等式, 并注意到对于 $d \geqslant 1$ 及 $c(d)>0$, 有 $\mathbf{P}\{X_{2k}=0\} \sim \frac{c(d)}{n^{d/2}}$, 就立即可以得到波利亚定理的结论: 对于 $d=1$ 和 $d=2$, 有

$$\mathbf{P}\{\sigma_1 < \infty\} = 1$$

(自返的概率为 1); 而对于 $d \geqslant 3$, 有

$$\mathbf{P}\{\sigma_1 < \infty\} < 1$$

(非自返的概率为正).

11. 我们来讨论关于区域 $C \subseteq E$ 中的泊松方程的狄利克雷问题 (参阅附录第 7 节 p339, 其中 E 为至多可数的集合: 寻找非负函数 $V=V(x)$, 使得

$$\mathbb{L}V(x) = -h(x), \quad x \in C,$$
$$V(x) = g(x), \quad x \in D = E \setminus C,$$

其中 $h(x)$ 与 $g(x)$ 为非负函数.

证明, 该问题的最小非负解 $V_D(x)$ 由下式所定义:

$$V_D(x) = \mathbf{E}_x\left[I(\tau(D)<\infty)g(X_{\tau(D)})\right] + I_C(x)\mathbf{E}_x\left[\sum_{k=0}^{\tau(D)-1} h(X_k)\right],$$

其中, $\tau(D) = \inf\{n \geqslant 0: X_n \in D\}$(如同往常, 若 $\{\cdot\}=\varnothing$, 则 $\tau(D)=\infty$).

提示: 写函数 $V_D(x) = \varphi_D(x) + \psi_D(x)$, 其中 $(\overline{D}=C)$:

$$\varphi_D(x) = \mathbf{E}_x\left[I(\tau(D)<\infty)g(X_{\tau(D)})\right],$$
$$\psi_D(x) = I_{\overline{D}}(x)\mathbf{E}_x\left[\sum_{k=0}^{\tau(D)-1} h(X_k)\right].$$

此二函数 $\varphi_D(x)$ 与 $\psi_D(x)$ 分别具有下述表达式:

$$\varphi_D(x) = I_D(x)g(x) + I_{\overline{D}}(x)T\varphi_D(x),$$
$$\psi_D(x) = I_{\overline{D}}(x)h(x) + I_{\overline{D}}(x)T\psi_D(x),$$

(其中 T 是一步转移算子; 参阅附录第 7 节 p339). 由这些等式可以推出

$$V_D(x) = I_D(x)g(x) + I_{\overline{D}}(x)[h(x) + TV_D(x)],$$

由此得知, 在区域 C 上, 这个函数是方程组的非负解, 有 $LV(x) = -h(x)$, 而对于 $x \in D$, 则有 $V(x) = g(x)$.

为对方程组的任意非负解 $V(x)$, 证明 $V(x) \geqslant V_D(x)$, 应当指出, 对于这些解, 有 $V(x) = I_D(x)g(x) + I_{\overline{D}}(x)[h(x)+TV(x)]$, 由该关系式得知

$$V(x) \geqslant I_D(x)g(x) + I_{\overline{D}}(x)h(x).$$

利用此不等式, 通过归纳法, 可证, 对任何 $n \geqslant 0$, 都有

$$V(x) \geqslant \sum_{k=0}^{n}(I_{\overline{D}}T^k)\left[I_D g + I_{\overline{D}}h\right](x).$$

由此即得

$$V(x) \geqslant \sum_{k\geqslant 0}(I_{\overline{D}}T^k)\left[I_D g + I_{\overline{D}}h\right](x) = \varphi_D(x) + \psi_D(x) = V_D(x).$$

12. 设 $X=(X_n)_{n\geqslant 0}$ 是 $\mathbb{Z}^d$ 中的简单对称随机游动, 而

$$\sigma(D) = \inf\{n>0: X_n \in D\}, \quad D \subseteq \mathbb{Z}^d,$$

并且 $\overline{D}$ 为有限集. 证明, 存在正常数 $c=c(D)$ 与 $\varepsilon=\varepsilon(D)<1$, 使得对一切 $x \in \overline{D}$, 都有

$$\mathbf{P}_x\{\sigma(D) \geqslant n\} < c\varepsilon^n.$$

(试比较第一章第 9 节中的不等式 (20).)

13. 在概率空间 $(\Omega, \mathscr{F}, \mathbf{P})$ 上考察两个分别自 x 与 y 开始的相互独立的简单对称随机游动 $X^1 = (X_n^1)_{n\geqslant 0}$ 与 $X^2 = (X_n^2)_{n\geqslant 0}$ $(X_0^1 = x,\ X_0^2 = y;\ x, y \in \mathbb{Z})$. 令 $\tau^1(x) = \inf\{n \geqslant 0 : X_n^1 = 0\}$, $\tau^2(y) = \inf\{n \geqslant 0 : X_n^2 = 0\}$. 试求概率 $\mathbf{P}\{\tau^1(x) < \tau^2(y)\}$.

14. 证明, 对于集合 $\mathbb{Z} = \{0, \pm 1, \pm 2, \cdots\}$ 上的简单对称随机游动 $X = (X_n)_{n\geqslant 0}$, $X_0 = 0$, 有

$$\mathbf{P}_0\{\tau(y) = N\} \sim \frac{|y|}{\sqrt{2\pi}} N^{-3/2}, \quad N \to \infty,$$

其中 $\tau(y) = \inf\{n \geqslant 0 : X_n = y\}$, $y \neq 0$. (试比较第二章第 4 节第 16 题中的断言.)

15. 我们来考察简单随机游动 $X = (X_n)_{n\geqslant 0}$, 其中 $X_n = x + \xi_1 + \cdots + \xi_n$, 而 $\xi_1, \xi_2, \cdots$ 是独立同分布的随机变量序列, 有 $\mathbf{P}\{\xi_1 = 1\} = p$, $\mathbf{P}\{\xi_1 = -1\} = q$, $p + q = 1$, 又 $x \in \mathbb{Z}$. 令 $\sigma(x) = \inf\{n > 0 : X_n = x\}$. 证明,

$$\mathbf{P}_x\{\sigma(x) < \infty\} = 2\min(p, q).$$

16. 考察上题中的随机游动, 令 $x = 0$. 以 $\mathscr{R}_n$ 表示序列 $X_0, X_1, \cdots, X_n\,(X_0 = 0)$ 中的不同的值的个数. 证明,

$$\mathbf{E}_0 \frac{\mathscr{R}_n}{n} \to |p - q|, \quad n \to \infty.$$

17. 设 $X = (X_n)_{n\geqslant 0}$ 是集合 $\mathbb{Z} = \{0, \pm 1, \pm 2, \cdots\}$ 上的简单随机游动, 其中 $X_0 = 0$, $X_n = \xi_1 + \cdots + \xi_n$, 而 $\xi_1, \xi_2, \cdots$ 是独立同分布的随机变量序列, 有 $\mathbf{P}\{\xi_1 = 1\} = p$, $\mathbf{P}\{\xi_1 = -1\} = q\ (= 1 - p)$, $0 < p < 1$. 证明, 序列 $|X| = (|X_n|)_{n\geqslant 0}$ 是以 $E = \{0, 1, 2, \cdots\}$ 为状态空间的马尔可夫链, 它的转移概率为:

$$p_{i,i+1} = \frac{p^{i+1} + q^{i+1}}{p^i + q^i} = 1 - p_{i,i-1}, \quad i > 0, \quad p_{01} = 1.$$

再证明, 对任何 $n \geqslant 1$ 及任何非负整数 $i,\ i_1,\ \cdots,\ i_{n-1}$, 都有

$$\mathbf{P}\left(X_n = i \mid |X_n| = i,\ |X_{n-1}| = i_{n-1},\ \cdots,\ |X_1| = i_1\right) = \frac{p^i}{p^i + q^i}.$$

18. 设 $\xi = (\xi_0, \xi_1, \xi_2, \cdots)$ 是独立同分布的随机变量序列, 有 $\mathbf{P}\{\xi_1 = 1\} = p$, $\mathbf{P}\{\xi_1 = -1\} = q$, $p + q = 1$. 定义 $X_n = \xi_n \xi_{n+1}$, $n \geqslant 0$. 试问, 序列 $X = (X_n)_{n\geqslant 0}$ 是否为马尔可夫链? 如果令 $Y_n = \frac{1}{2}(\xi_{n-1} + \xi_n)$, $n \geqslant 1$, 那么 $Y = (Y_n)_{n\geqslant 1}$ 是否为马尔可夫链?

19. 设 $X = (X_n)_{n\geqslant 0}$ 是 $\mathbb{Z} = \{0, \pm 1, \pm 2, \cdots\}$ 上的简单对称随机游动. 令 $\sigma_1, \sigma_2, \cdots$ 是随机游动 X 回归状态 0 的时刻 (其中 $X_0 = 0$). (即有 $\sigma_1 = \inf\{n > 0 : X_n = 0\}$, $\sigma_2 = \inf\{n > \sigma_1 : X_n = 0\}$, 如此等等.)

证明,

(a) $\mathbf{P}_0\{\sigma_1 < \infty\} = 1$.

(b) $\mathbf{P}_0\{X_{2n} = 0\} = \sum_{k\geqslant 1} \mathbf{P}_0\{\sigma_k = 2n\}$.

(c) 矩母函数 $\mathbf{E}_0\, z^{\sigma_k} = (\mathbf{E}_0\, z^{\sigma_1})^k$ $(|z| < 1)$.

(d) $\sum_{n\geqslant 0} \mathbf{P}_0\{X_{2n} = 0\} z^{2n} = \dfrac{1}{1 - \mathbf{E}_0\, z^{\sigma_1}}$.

(e) 母函数 $\mathbf{E}_0\, z^{\sigma_1} = 1 - \sqrt{1 - z^2}$, 因此, 结合 (d), 得知 $\sum_{n\geqslant 0} \mathbf{P}_0\{X_{2n} = 0\} z^{2n} = \dfrac{1}{\sqrt{1 - z^2}}$.

(f) $\mathbf{E}N(k) = 1$, 其中 k 是 $\mathbb{Z}$ 中的任一状态, 但 $k \neq 0$, 而 $N(k)$ 是随机游动在第一次回归状态 0 (即时刻 σ_1) 之前所访问状态 k 的次数.

20. 设 $X = (X_n)_{n\geqslant 0}$ 与 $Y = (Y_n)_{n\geqslant 0}$ 是 $\mathbb{Z}^d$ 上的两个简单对称随机游动, $d \geqslant 1$. 令

$$R_n = \sum_{i=0}^{n} \sum_{j=0}^{n} I(X_i = X_j).$$

证明, 在 $d = 1$ 时, 数学期望 $\mathbf{E}R_n$ 是两个游动在时刻 n 时相交的平均次数的一个数字特征 (包括多次经过同一个点的次数), 对于足够大的 n, 它的阶为 $cn^{3/2}$, 其中常数 $\mathrm{c} > 0$. (现已知道, 例如可以参阅 [75], 对于不同的 $d > 1$, 当 n 足够大时, 分别有

$$\mathbf{E}R_n = \begin{cases} \mathrm{c}n, & d = 2; \\ \mathrm{c}\sqrt{n}, & d = 3; \\ \mathrm{c}\ln n, & d = 4; \\ \mathrm{c}, & d \geqslant 5, \end{cases}$$

其中, $\mathrm{c} = c_d$ 是与 d 有关的常数. 请尝试证明所列举的关于 $\mathbf{E}R_n$ 的各个渐近表达式, 并求出常数 c_d 的值.)

21. 设 B 是 $\mathbb{Z}^d$ 中的有限集, $f = f(x)$ 为定义在 $B \cup \partial B$ 上的函数, 其中 $\partial B = \{x \notin B : \|x - y\| = 1$ 对某个 $y \in B\}$. 假设函数 $f = f(x)$ 在 B 中上调和 (即 $Tf(x) \geqslant f(x)$, $x \in B$, 其中 T 为一步转移算子; 参阅附录第 7 节 p339). 证明, 这样的函数满足最大值原理:

$$\sup_{x\in B\cup\partial B} f(x) = \sup_{x\in\partial B} f(x).$$

22. 证明, 一切在 $\mathbb{Z}^d$ 上调和的有界函数都是常数.

23. 设 $C \subset \mathbb{Z}^d$, $g = g(x)$ 为 ∂C 上的有界函数. 证明, 问题

$$\Delta V(x) = 0, \quad x \in C, \quad V(x) = h(x), \quad x \in \partial C$$

的一切有界解 $V = V(x)$ 均满足如下等式:

$$V_{\partial C}(x) = \mathbf{E}_x [g(X_{\tau(\partial C)}) I(\tau(\partial C) < \infty)] + \alpha \mathbf{P}\{\tau(\partial C) = \infty\},$$

其中, $\alpha \in \mathbb{R}$, $\tau(\partial C) = \inf\{n \geqslant 0: X_n \in \partial C\}$.

24. 所谓齐次狄利克雷问题, 就是寻找函数 $V = V(x)$, 使得它在区域 $C \subseteq \mathbb{Z}^d$ 上调和 $(\Delta V(x) = 0,\ x \in C)$, 在 ∂C 上, 有 $V(x) = g(x)$, 其中 $g = g(x)$ 是给定的函数. 证明关于该问题解的如下结果:

(a) 若 $d \leqslant 2$ 且函数 $g = g(x)$ 有界, 则在有界函数类中, 解存在、唯一, 且满足等式 $V_{\partial C}(x) = \mathbf{E}_x g(X_{\tau(\partial C)})$.

(b) 若 $d \geqslant 3$, 函数 $g = g(x)$ 有界, 且对任何 $x \in C$, 都有 $\mathbf{P}\{\tau(\partial C) < \infty\} = 1$, 则在有界函数类中, 解存在、唯一, 且满足 (a) 中的等式.

25. 设 $X = (X_n)_{n \geqslant 0}$ 是 $\mathbb{Z}^d$ 上的简单对称随机游动, $d \geqslant 1$. 证明, 如果区域 $C \subset \mathbb{Z}^d$ 有界, 并且其边界 ∂C 为集合 $\{x \in \mathbb{Z}^d:\ x \notin C,\ \|x - y\| = 1 \text{ 对某个 } y \in C\}$, 则狄利克雷问题:

寻求 $C \cup \partial C$ 上的函数 $V = V(x)$, 使之满足

$$\Delta V(x) = -h(x), \quad x \in C,$$
$$V(x) = g(x), \quad x \in \partial C,$$

就有由下式所给出的唯一解:

$$V_{\partial C}(x) = \mathbf{E}_x g(X_{\tau(\partial C)}) + \mathbf{E}_x \left[\sum_{k=0}^{\tau(\partial C)-1} h(X_k)\right], \quad \text{其中 } \mathbf{E}_x \left[\sum_{k=0}^{\tau(\partial C)-1} h(X_k)\right] < \infty.$$

提示: 利用第 11 题提示中所述的方法, 并注意, 关于区域 C 的有界性假设蕴涵等式 $\mathbf{P}_x\{\tau(\partial C) < \infty\} = 1,\ x \in C$ (比照第 11 题).

26. 设 $S = (S_n)_{n \geqslant 0}$ 是 $\mathbb{Z} = \{0, \pm 1, \pm 2, \cdots\}$ 上的简单对称随机游动, 有 $S_0 = 0$, $S_n = \xi_1 + \cdots + \xi_n$, 其中 $\xi_1, \xi_2, \cdots$ 是相互独立的伯努利随机变量, $\mathbf{P}\{\xi_i = 1\} = \mathbf{P}\{\xi_i = -1\} = 1/2$. 令 $\tau = \inf\{n \geqslant 0 : S_n = 1\}$. 第七章第 2 节第 18 题曾经运用鞅方法, 证明了, 对于 $\alpha \leqslant 1$, 有

$$\mathbf{E}\alpha^{\tau} = \alpha^{-1}[1 - \sqrt{1 - \alpha^2}].$$

试以严马尔可夫性为基础, 证明, 函数 $\varphi(\alpha) = \mathbf{E}\alpha^{\tau}$ 满足关系式 $\varphi(\alpha) = \frac{1}{2}\alpha + \frac{1}{2}\alpha\varphi^2(\alpha)$, 并由此推出关于母函数 $\varphi(\alpha) = \mathbf{E}\alpha^{\tau}$ 的表达式.

27. 本题涉及由 T. 埃伦弗斯特和 P. 埃伦弗斯特所提出的模型中的若干计算问题, 该模型是用来解释玻尔兹曼热扩散动力学理论中的 "不可逆性" 与 "自返性" 之间的并不存在的矛盾的.

如所周知, 产生于分子结构, 并将热交换视为扩散随机过程的这一理论是由玻尔兹曼为解释热力学理论而提出来的 (颇具非凡特色), 其中采纳了这样一个假设, 即热扩散以不可逆的方式朝着热平衡的方向进行. 尽管玻尔兹曼认为, 在系统的状态中进行着热平衡的建立, 并且朝着最大熵的状态运动, 但在他所倡导的 "随机" 理论中并未排除, 至少在理论上没有排除, 这样的一种可能性, 即随着时间的延续, 系统可以回到自己最初的热力学非平衡态, 而这正是人们对动力学理论进行批评的一个基本点. (庞加莱曾经对描述保测变换的动力系统指出过 "自返性" 的可能性; 参阅第五章第 1 节.)

玻尔兹曼自己断言, 在 "不可逆性" 与实际上是观察不到的 "自返性" 之间并不存在矛盾, 因为在随机系统中, 返回到肉眼可见的非平衡态是根本做不到的, 而要发生这种情况, 只有经过如此之长的时间, 以致于我们根本不能见到这种状态.

从物理学的观点来看, 由两位埃伦弗斯特所建立的以马尔可夫链的语言所表述的模型, 在刻画两个与外界环境隔绝的物体之间的热交换方面是非常成功的. 它给出了由一个状态过渡到另一个状态的平均时间的审慎而有趣的数量分析.

设 $E = \{0, 1, \cdots, 2k\}$, 其中状态 i 解释为 "在箱子 A 中有 i 个分子"(关于埃伦弗斯特模型的更加详细的介绍, 可以参阅《概率》第二卷 p293). 分别以

$$\tau(i) = \inf\{n \geqslant 0 : X_n = i\}, \quad \sigma(i) = \inf\{n > 0 : X_n = i\}$$

表示首次进入状态 i 或首次回到状态 i 的时刻, 若 $X_0 = i$(如同往常, $\inf \varnothing = \infty$). 证明,

(a) $\mathbf{E}_i\, \sigma(i) = 2^{2k} \frac{i!(2k-i)!}{(2k)!}$, 特别地, 返回状态 0 的平均时间为 $\mathbf{E}_0\, \sigma(0) = 2^{2k}$.

(b) $\mathbf{E}_k\, \tau(0) = \frac{1}{2k} 2^{2k}(1 + O(k))$.

(c) $\mathbf{E}_0\, \tau(k) = k \ln k + k + O(1)$. (在 [19] 中有如下的数值计算: 如果 $k = 10000$, 并且分子间的交换每秒钟进行一次, 则平均时间 $\mathbf{E}_0\, \tau(k)$ 少于 29 小时, 然而, 与此同时, $\mathbf{E}_k\, \tau(0)$ 却是一个天文数字: 10^{6000} 年 (!).)

§9. 马尔可夫链的最优停止问题

1. 试举例说明, 对于具有可数状态空间的马尔可夫链, 有可能 (在类 $\mathfrak{M}_0^{\infty}$ 中) 不存在最优停时.

2. 设 τ_y 如定理 2 的证明中所定义. 证明, τ_y 是马尔可夫时.

3. 证明, 第 7 小节在 "任性的未婚妻问题" 中所引入的序列 $X = (X_1, X_2, \cdots)$ 形成齐次马尔可夫链.

4. 设 $X = (X_n)_{n \geqslant 0}$ 为取值于 $\mathbb{R}$ 的齐次马尔可夫链, 具有转移函数 $P = P(x; B)$,

$x \in \mathbb{R}$, $B \in \mathscr{B}(\mathbb{R})$. 我们说, $\overline{\mathbb{R}}$ 值的函数 $f = f(x)$, $x \in \mathbb{R}$ 是调和的 (或者为 P-调和的, 或者关于同 P 之比为调和的), 如果有

$$\mathbf{E}_x |f(X_1)| = \int_{\mathbb{R}} |f(y)| P(x; dy), \quad x \in \mathbb{R}$$

与

$$f(x) = \int_{\mathbb{R}} f(y) P(x; dy), \quad x \in \mathbb{R}. \qquad (*)$$

如果在式 $(*)$ 中将等号 "=" 换为不等号 "$\geqslant$", 则说 f 是上调和的 (参阅附录第 7 节).

证明, 如果 f 是上调和函数, 则对于所有的 $x \in \mathbb{R}$, 序列 $(f(X_n))_{n\geqslant 0}$ (其中 $X_0 = 0$) 都是上鞅 (关于同测度 $\mathbf{P}_x$ 的比值).

5. 证明, (40) 式中的时刻 τ 属于 $\mathfrak{M}_1^\infty$ 类.

6. 对于第 8 节中的各种简单随机游动的例子, 仿照第 6 小节例 1, 讨论最优停时问题:

$$S_N(x) = \sup_{\tau \in \mathfrak{M}_0^N} \mathbf{E}_x g(X_\tau)$$

与

$$S(x) = \sup_{\tau \in \mathfrak{M}_0^\infty} \mathbf{E}_x g(X_\tau).$$

7. (受控马尔可夫链与优化问题) 给定一个转移概率矩阵的族 $\{\mathbb{P}(a),\ a \in A\}$, 其中 $\mathbb{P}(a) = \|p_{ij}(a)\|$ 依赖于参数 $a \in A$, 而 A 是某个可视为所有可能 "控制值" 的返回值集合. 相空间 $E = \{i, j, \cdots\}$ 有限或可数. (所有取值于 A 的函数 $u = u(i)$, $i \in \mathbb{R}$ 均被理解为在状态 $i \in E$ 下的可允许的 "控制".)

对于每个被选的控制 $u = u(i)$, $i \in E$, 我们用 $\mathbb{P}^u$ 表示转移概率矩阵 $\|p_{ij}(u(i))\|$, 根据其构造 (例如, 按照第二章第 9 节中的约内斯库-图尔其定理) 相应的序列空间 E^∞ 中的概率分布族 $\mathbf{P}_i^u$, $i \in E$, 将其视为由状态 i 出发 (即 $X_0 = i$) 的 (被 "控制"u 所控制的) 受控马尔可夫链 $X = (X_n)_{n\geqslant 0}$ 的概率分布.

设 C 是 E 中的某个区域, $D = E \backslash C$. 假设在 C 中给定了函数 $h = h(i, a)$, $i \in C$, $a \in A$ 与 $g = g(i, a)$, $i \in D$, $a \in A$, 并且它们都是非负值函数.

如果控制 $u = u(i)$, $i \in E$ 被选取出来, 则令 $h^u(i) = h(i, u(i))$ 与 $g^u(i) = g(i, u(i))$.

与每个控制 $u = u(i)$, $i \in E$ 相关联的在状态 $j \in E$ 下的 "收入" 为:

$$V^u(j) = \mathbf{E}_j^u \left[g^u(X_{\tau(D)}) I(\tau(D) < \infty) + \sum_{k=0}^{\tau(D)-1} h^u(X_k) \right],$$

其中 $\tau(D) = \inf\{n \geqslant 0: X_n \in D\}$. ($V^u(j)$ 的含义是清楚的: 这就是若选取开始于状态 $X_0 = j$ 的控制 $u = u(i)$, $i \in E$, 所可以赢得的钱数; 值 $g^u(i)$ 刻画了理论上的钱数, 而 $h^u(i)$ 则是在状态 i 下的现在的钱数.)

对于所考察的受控马尔可夫链的最优化问题就是寻找 "价格"

$$V^*(j) = \sup_u V^u(j), \quad j \in E$$

和最优控制 $u^* = u^*(i)$, $i \in E$, 如果它们存在.

证明如下的所谓检验定理: 假设

(i) 存在函数 $V = V(j)$, $j \in E$, 使得

$$V(j) = \sup_{a \in A} \left\{ \sum_{j \in E} p_{ji}(a) V(i) + h(j, a) \right\}, \quad j \in C,$$

与

$$V(j) = \sup_{a \in A} g(j, a), \quad j \in D.$$

(ii) 在所考察的控制类中, 存在控制 $u^* = u^*(i)$, $i \in E$, 使得对于每个固定的 j, 上述二式中的上确界均在 $a = u^*(j)$ 处达到.

则控制 $u^* = u^*(i)$, $i \in E$, 是最优的: 对任何可允许的控制 u, 都有

$$V^{u^*}(j) \geqslant V^u(j) \quad 且 \quad V^{u^*}(j) = V(j),\ j \in E.$$

提示: 利用事实: 对任何控制 u, 都有

$$V(j) \geqslant T^u V(j) + h^u(j), \quad j \in C, \quad 其中 \quad T^u V(j) = \mathbf{E}_j^u V(X_1),$$

与

$$V(j) \geqslant g^u(j), \quad j \in D,$$

而对于控制 u^*, 则上述二式均成立等号.

8. (无序问题) 在第六章第 7 节第 8 题中, 曾经引入过贝叶斯风险:

$$R^\pi(\tau) = \mathbf{P}^\pi\{\tau < \theta\} + c\mathbf{E}^\pi(\tau - \theta)^+,$$

并且指出, $\inf R^\pi$(其中的下确界系对所有 $\mathbf{P}^\pi$-有限的马尔可夫时 τ 的类 $\mathfrak{M}_0^\infty$ 所取, $(\pi \in [0,1])$) 在如下的时刻达到:

$$\tau^* = \inf\{n \geqslant 0: \pi_n \geqslant A\}, \qquad (*)$$

其中 A 是依赖于 c 与 p 的常数.

证明,

(a) 贝叶斯风险 $R^\pi(\tau)$ 具有如下表达式:

$$R^\pi(\tau) = \mathbf{E}^\pi \left\{ (1-\pi_\tau) + cI(\tau \geqslant 1) \sum_{k=0}^{\tau-1} \pi_k \right\}.$$

(b) 在关于马尔可夫链 $(\pi_n)_{n\geqslant 0}$ 的最优停时问题

$$R^\pi(\tau) = \inf_{\tau \in \mathfrak{M}_0^\infty} \mathbf{E}^\pi \left\{ (1-\pi_\tau) + cI(\tau \geqslant 1) \sum_{k=0}^{\tau-1} \pi_k \right\}$$

中, 由 (*) 所确定的停时 τ^* 是最优的.

附录　本书所用到的组合论与概率论中的基本符号与重要概念简介

§1. 组合论基础

"概率计算" 从一开始就与组合学密切联系在一起, 因为它需要用组合方法计算使得所考察的事件发生的可能情况的数目. 这些方法即使在现代概率论中也依然占据着显著的地位, 尤其是在那些具有有限个试验结果的场合 (《初等概率论》). 组合方法构成了离散数学, 包括它的各个分支, 诸如图论, 算法理论等等中的最本质的部分.

我们来大致介绍一下组合学的某些基本概念及其结果, 这些内容曾经在《概率》第一、二卷的基本内容以及本习题集中被运用到.

- 设 $\boldsymbol{A}$ 是具有有限个元素 $a_1, \cdots, a_N$ 的组 ($|\boldsymbol{A}| = N$). 如果这些元素各不相同, 那么就将 $\boldsymbol{A}$ 称为集合, 并记作

$$\boldsymbol{A} = \{a_1, \cdots, a_N\}.$$

在这里, 元素的书写顺序是无关紧要的, 重要的是其中究竟有哪些成员. 因而, 组 $\{1,2,3\}$ 与组 $\{2,3,1\}$ 所确定的是同一个由 "点"$\{1\},\{2\}$ 与 $\{3\}$ 所构成的集合.

与每个集合 $\boldsymbol{A} = \{a_1, \cdots, a_N\}$ 联系着两种不同类型的大小为 n 的样本 (有时也称为长度为 n 的序列):

$$(a_{i_1}, \cdots, a_{i_n}) \quad 与 \quad [a_{i_1}, \cdots, a_{i_n}],$$

其中 $i_1,\cdots,i_n\in\{1,\cdots,N\}$, 而元素 $a_{i_j}\in\boldsymbol{A}$, 并且对于不同的 j, a_{i_j} 有可能彼此相同.

用符号 $(a_{i_1},\cdots,a_{i_n})$ 表示*有序样本*, 即元素的排列顺序是关键因素的样本.

用符号 $[a_{i_1},\cdots,a_{i_n}]$ 表示*无序样本*, 即只关心其中的成员有哪些, 而不计较它们的排列顺序的样本.

如此一来, 样本 (a_4,a_1,a_3,a_1) 与样本 (a_1,a_1,a_4,a_3) 尽管组成成员相同, 但排列顺序不同, 因而是互不相同的. 而 $[a_4,a_1,a_3,a_1]$ 与 $[a_1,a_1,a_4,a_3]$ 则视为同一个样本.

如果 $(\cdots)$ 与 $[\cdots]$ 是由 $\boldsymbol{A}$ 按 "无放回抽样" 所形成的样本, 则它们中的所有成员各不相同, 因而 $n\leqslant N$.

在 "有放回抽样" 的情况下, $(\cdots)$ 与 $[\cdots]$ 中的成员可以重复出现, 因而 n 可以任意大.

设 $\boldsymbol{A}$ 为集合, 有 $|\boldsymbol{A}|=N$. 如果子集 $D_i\subseteq\boldsymbol{A}$, $1\leqslant i\leqslant n$, 满足条件 $D_i\neq\varnothing$, $D_i\cap D_j=\varnothing$, $i\neq j$ 且 $D_1+\cdots+D_n=\boldsymbol{A}$, 则将 $\mathscr{D}=\{D_1,\cdots,D_n\}$ 称为 $\boldsymbol{A}$ 的*一个分割*, $n\leqslant N$. 集合 D_i 称为分割 $\mathscr{D}$ 的*原子*或*分割类*.

- $\boldsymbol{A}=\{a_1,\cdots,a_N\}$ 的元素构成的样本数目的计算.

组合计数	**意义**
(a) $(N)_n\equiv N(N-1)\cdots(N-n+1)$ (由 N 个元素中选出 n 个的排列数目, $1\leqslant n\leqslant N$)	$\boldsymbol{A}$ 为集合, 有 $|\boldsymbol{A}|=N$. 由 $\boldsymbol{A}$ 中的元素按 "无放回抽样" 所形成的大小为 n 的有序样本 $(\cdots)$ 的数目
(b) $C_N^n\equiv\frac{(N)_n}{n!}\ \left(=\frac{N!}{n!(N-n)!}\right)$ (由 N 个元素中选出 n 个的组合数目, 二项式系数)	$\boldsymbol{A}$ 为集合, 有 $|\boldsymbol{A}|=N$. 由 $\boldsymbol{A}$ 中的元素按 "无放回抽样" 所形成的大小为 n 的无序样本 $[\cdots]$ 的数目
(c) N^n	$\boldsymbol{A}$ 为集合, 有 $|\boldsymbol{A}|=N$. 由 $\boldsymbol{A}$ 中的元素按 "有放回抽样" 所形成的大小为 n 的有序样本 $(\cdots)$ 的数目
(d) C_{N+n-1}^n	$\boldsymbol{A}$ 为集合, 有 $|\boldsymbol{A}|=N$. 由 $\boldsymbol{A}$ 中的元素按 "有放回抽样" 所形成的大小为 n 的无序样本 $[\cdots]$ 的数目

关于上列各组合数的其他解释可以参阅第一章第 1 节与第 2 节中的习题. 特别是, 根据第 1 节第 3 题, "由 n 个 1 与 $N-n$ 个 0 所构成的大小为 N 的有序样本 $(\cdots)$ 的数目等于 C_N^n". 该结果与二项分布有关, 在初等概率论中非常重要.

例 1. 设 $\boldsymbol{A}=\{a_1,a_2,a_3,a_4\}$, $|\boldsymbol{A}|=4$, $n=2$.

(a) $(4)_2=4(4-1)=12$, 相应的 12 个有序样本是:

$$(a_1,a_2),\quad(a_1,a_3),\quad(a_1,a_4),\quad(a_2,a_1),\quad(a_2,a_3),\quad(a_2,a_4),$$
$$(a_3,a_1),\quad(a_3,a_2),\quad(a_3,a_4),\quad(a_4,a_1),\quad(a_4,a_2),\quad(a_4,a_3).$$

(b) $C_4^2=\frac{4!}{2!2!}=6$, 相应的 6 个无序样本是:

$$[a_1,a_2],\quad[a_1,a_3],\quad[a_1,a_4],\quad[a_2,a_3],\quad[a_2,a_4],\quad[a_3,a_4].$$

(c) $4^2=16$, 在 (a) 中的 12 个有序样本的基础上, 再添加 4 个样本:

$$(a_1,a_1),\quad(a_2,a_2),\quad(a_3,a_3),\quad(a_4,a_4).$$

(d) $C_{4+2-1}^2=C_5^2=\frac{5!}{3!2!}=10$, 在 (b) 中的 6 个无序样本的基础上, 再添加 4 个样本: $[a_1,a_1]$, $[a_2,a_2]$, $[a_3,a_3]$, $[a_4,a_4]$.

- $\boldsymbol{A}=\{a_1,\cdots,a_N\}$ 的子集数目与分割数目的计算.

组合计数	**意义**
(e) 2^N	$\boldsymbol{A}$ 为集合, 有 $|\boldsymbol{A}|=N$. $\boldsymbol{A}$ 中的所有不同子集的数目 (包括空集 $\varnothing$ 与集合 $\boldsymbol{A}$ 自身)
(f) $C_N^n=\frac{N!}{n!(N-n)!}$	$\boldsymbol{A}$ 为集合, 有 $|\boldsymbol{A}|=N$. $\boldsymbol{A}$ 中的由 n 个元素构成的不同子集 $\mathbf{D}\subseteq\boldsymbol{A}$ 的数目 ($|D|=n$, $0\leqslant n\leqslant N$; 当 $n=0$ 时, 有 $\mathbf{D}=\{\varnothing\}$, $C_N^0=1$)
(g) $C_N(n_1,\cdots,n_r)=\frac{N!}{n_1!\cdots n_r!}$ ("多组组合" 或 "多项式系数" $n_1+\cdots+n_r=N$)	$\boldsymbol{A}$ 为集合, 有 $|\boldsymbol{A}|=N$. 将 $\boldsymbol{A}$ 划分为 r 个互不相交的子集 $D_1,\cdots,D_r$, 其中分别有 $n_1,\cdots,n_r$ 个元素($n_1+\cdots+n_r=N$) 的所有不同分割 $\mathscr{D}=\{D_1,\cdots,D_r\}$ 的数目, $|D_1|=n_1,\cdots,|D_r|=n_r$, $r\leqslant n$
(h) $D_N(\lambda_1,\cdots,\lambda_N)=\frac{N!}{(1!)^{\lambda_1}\cdots(N!)^{\lambda_N}(\lambda_1)!\cdots(\lambda_N)!}$ 所有 ($\lambda_i\geqslant0$, $\sum\limits_{i=1}^N i\lambda_i=N$)	$\boldsymbol{A}$ 为集合, 有 $|\boldsymbol{A}|=N$. 对 $\boldsymbol{A}$ 所作的如下形式的不同 "分块分割" 的数目: $\mathscr{D}=\{D_{11},\cdots,D_{1\lambda_1};\cdots;D_{N1},\cdots,D_{N\lambda_N}\}$, 其中 "块" $[\![D_{i1},\cdots,D_{i\lambda_i}]\!]$ 由 λ_i 个 i 元子集组成 ($|D_{ik}|=i$, $1\leqslant k\leqslant\lambda_i$); 若 $\lambda_i=0$, 则相应的 "块" 无定义, 亦即 $\mathscr{D}$ 中没有这样的 "块"
(i) $S_N^n=\sum D_N(\lambda_1,\cdots,\lambda_N)$ (其中的求和系对所有满足如下条件的 $\lambda_1,\cdots,\lambda_N$ 进行: $\lambda_i\geqslant0$, $\sum\limits_{i=1}^N\lambda_i=n$, $\sum\limits_{i=1}^N i\lambda_i=N$)	$\boldsymbol{A}$ 为集合, 有 $|\boldsymbol{A}|=N$. $\boldsymbol{A}$ 的所有可能的恰含有 n 个类的分割 $\mathscr{D}$ 的数目

$S_N^n (1 \leqslant n \leqslant N)$ 被称为二阶斯特林数.

(j) $B_N = \sum\limits_{n=1}^{N} S_N^n$ $\qquad$ $\boldsymbol{A}$ 为集合, 有 $|\boldsymbol{A}| = N$. $\boldsymbol{A}$ 的所有可能分割的数目 B_N 被称为贝尔数

(关于 (f) ~ (j) 中所提到的各种组合数, 可参阅第一章第 2 节中的习题①.)

例 2. 仍设 $\boldsymbol{A} = \{a_1, a_2, a_3, a_4\}$, $|\boldsymbol{A}| = 4$, $n = 2$.

(e) $2^4 = 16$. 这 16 个子集是:

$$\varnothing, \quad \{a_1\}, \quad \{a_2\}, \quad \{a_3\}, \quad \{a_4\}$$
$$\{a_1, a_2\}, \quad \{a_1, a_3\}, \quad \{a_1, a_4\}, \quad \{a_2, a_3\}, \quad \{a_2, a_4\}, \quad \{a_3, a_4\}$$
$$\{a_1, a_2, a_3\} \quad \{a_1, a_2, a_4\} \quad \{a_1, a_3, a_4\} \quad \{a_2, a_3, a_4\} \qquad \{a_1, a_2, a_3, a_4\}.$$

(f) $C_4^2 = 6$. 相应的 6 个 2 元子集是:

$$\{a_1, a_2\}, \quad \{a_1, a_3\}, \quad \{a_1, a_4\}, \quad \{a_2, a_3\}, \quad \{a_2, a_4\}, \quad \{a_3, a_4\}.$$

(g) 设 $r = 2$, $n_1 = 1$, $n_2 = 3$, 于是 $C_4(1, 3) = \frac{4!}{1!3!} = 4$. 相应的 4 种分割为

$$\{a_1\} \text{ 与 } \{a_2, a_3, a_4\}, \qquad \{a_2\} \text{ 与 } \{a_1, a_3, a_4\},$$
$$\{a_3\} \text{ 与 } \{a_1, a_2, a_4\}, \qquad \{a_4\} \text{ 与 } \{a_1, a_2, a_3\}.$$

(h) $\lambda_1 = 2$, $\lambda_2 = 1$, $\lambda_3 = 0$, $\lambda_4 = 0$, $\sum\limits_{i=1}^{4} i\lambda_i = 4$,

$$D_4(1, 2, 0, 0) = \frac{4!}{(1!)^2(2!)^1(3!)^0(4!)^0 2!1!0!0!} = 6.$$

相应的 "分块分割" 是:②

$$[\![\{a_1\}, \{a_2\}]\!] \text{ 与 } [\![\{a_3, a_4\}]\!], \qquad [\![\{a_1\}, \{a_3\}]\!] \text{ 与 } [\![\{a_2, a_4\}]\!],$$
$$[\![\{a_1\}, \{a_4\}]\!] \text{ 与 } [\![\{a_2, a_3\}]\!], \qquad [\![\{a_2\}, \{a_3\}]\!] \text{ 与 } [\![\{a_1, a_4\}]\!],$$
$$[\![\{a_2\}, \{a_4\}]\!] \text{ 与 } [\![\{a_1, a_3\}]\!], \qquad [\![\{a_3\}, \{a_4\}]\!] \text{ 与 } [\![\{a_1, a_2\}]\!].$$

①关于一阶斯特林数可参阅本附录第 3 节 —— 译者注.

②原书误写为:

$$[\![\{a_1\}, \{a_2\}]\!] \text{ 与 } [\![\{a_3\}, \{a_4\}]\!], \qquad [\![\{a_1\}, \{a_3\}]\!] \text{ 与 } [\![\{a_2\}, \{a_4\}]\!],$$
$$[\![\{a_1\}, \{a_4\}]\!] \text{ 与 } [\![\{a_2\}, \{a_3\}]\!], \qquad [\![\{a_2\}, \{a_3\}]\!] \text{ 与 } [\![\{a_1\}, \{a_4\}]\!],$$
$$[\![\{a_2\}, \{a_4\}]\!] \text{ 与 } [\![\{a_1\}, \{a_3\}]\!], \qquad [\![\{a_3\}, \{a_4\}]\!] \text{ 与 } [\![\{a_1\}, \{a_2\}]\!]$$

—— 译者注.

(i) $S_4^2 = D_4(0, 2, 0, 0) + D_4(1, 0, 1, 0) = 4 + 3 = 7$. 相应的分割是:

$$\{a_1\} \text{ 与 } \{a_2, a_3, a_4\}, \quad \{a_2\} \text{ 与 } \{a_1, a_3, a_4\}, \quad \{a_3\} \text{ 与 } \{a_1, a_2, a_4\},$$
$$\{a_4\} \text{ 与 } \{a_1, a_2, a_3\}, \quad \{a_1, a_2\} \text{ 与 } \{a_3, a_4\}, \quad \{a_1, a_3\} \text{ 与 } \{a_2, a_4\},$$
$$\{a_1, a_4\} \text{ 与 } \{a_2, a_3\}.$$

可类似地算得:

$$S_4^1 = 1, \quad S_4^2 = 7, \quad S_4^3 = 6, \quad S_4^4 = 1,$$
$$S_5^1 = 1, \quad S_5^2 = 15, \quad S_5^3 = 25, \quad S_5^4 = 10, \quad S_5^5 = 1,$$
$$S_6^1 = 1, \quad S_6^2 = 31, \quad S_6^3 = 90, \quad S_6^4 = 65, \quad S_6^5 = 15, \quad S_6^6 = 1.$$

(j) 由 (i) 可知

$$B_4 = S_4^1 + S_4^2 + S_4^3 + S_4^4 = 1 + 7 + 6 + 1 = 15,$$
$$B_5 = 1 + 15 + 25 + 10 + 1 = 52,$$
$$B_4 = 1 + 31 + 90 + 65 + 15 + 1 = 203.$$

$N = 4$ 时的 15 个相应的分割为:

$$\{a_1, a_2, a_3, a_4\}; \qquad \{a_1\}, \{a_2\}, \{a_3\}, \{a_4\};$$
$$\{a_1\} \text{ 与 } \{a_2, a_3, a_4\}, \qquad \{a_2\} \text{ 与 } \{a_1, a_3, a_4\},$$
$$\{a_3\} \text{ 与 } \{a_1, a_2, a_4\}, \qquad \{a_4\} \text{ 与 } \{a_1, a_2, a_3\}$$
$$\{a_1, a_2\} \text{ 与 } \{a_3, a_4\}, \qquad \{a_1, a_3\} \text{ 与 } \{a_2, a_4\}, \qquad \{a_1, a_4\} \text{ 与 } \{a_2, a_3\},$$
$$\{a_1\}, \{a_2\} \text{ 与 } \{a_3, a_4\}, \qquad \{a_1\}, \{a_3\} \text{ 与 } \{a_2, a_4\}, \qquad \{a_1\}, \{a_4\} \text{ 与 } \{a_2, a_3\},$$
$$\{a_2\}, \{a_3\} \text{ 与 } \{a_1, a_4\}, \qquad \{a_2\}, \{a_4\} \text{ 与 } \{a_1, a_3\}, \qquad \{a_3\}, \{a_4\} \text{ 与 } \{a_1, a_2\}.$$

- 组合学方法的重要作用不仅体现在与初等概率论有关的组合问题的计数之中. 它们还可以成功地用来建立各种关系式, 其中的一个例子就是建立如下的等式:

$$n^N = \sum_{k=1}^{n} S_N^k (n)_k,$$

其中 $1 \leqslant n \leqslant N$, 而 S_N^k 是二阶斯特林数. (注意: $S_N^1 = S_N^N = 1$, 并假定 $S_N^0 = 0$, $S_N^n = 0$, $n > N$.)

证明这类关系式的组合论的方法的思想是: 其左端与右端表述同一种可能性的数目, 但两端是通过不同的计算方法得到的.

以上面的等式为例, 我们来说明其建立过程.

设 A 与 B 是两个集合, 它们中的元素个数分别为 $|A| = N$, $|B| = n$. 我们来考察函数 $y = f(x)$, 它对所有 $x \in A$ 有定义, 而取值于 $y \in B$. 由于对于 A 中的所有 N 个元素 x, 都可以让其对应于 B 中的 n 个元素 y 中的每一个, 所以在 A 与 B 共可建立 n^N 种不同的映射.

现在来用另一种方法计算各种不同的函数的可能数目. 对于每个 $y \in B$, 我们考察相应的逆像 $f^{-1}(y) = \{x: \ y = f(x)\}$. 选取某个集合 $C \subseteq B$, 有 $|C| = k$, $1 \leqslant k \leqslant n$. 设函数 $y = f(x)$ 的值域恰好就是集合 C. 由于 $|C| = k$, 所以集合 A 被划分为 k 个互不相交的非空子集, 在每个子集上, 函数各取一个值. 共有 S_N^k 种不同的这种分割, 而在各个子集上的取值则有 $(k)_k = k!$ 种不同的安排方式. 所以值域恰好就是集合 C 的函数 $y = f(x)$ 共有 $S_N^k\, k!$ 个.

选择 B 的 k 元子集 C 的方法有 C_n^k 种. 所以, 若 $|A| = N$, $|B| = n$, 则对所有 $x \in A$ 有定义, 并取值于 $y \in B$ 的函数 $y = f(x)$ 的数目就是

$$\sum_{k=1}^{n} C_n^k S_N^k\, k! = \sum_{k=1}^{n} S_N^k (n)_k.$$

由于按照第一种方法计算时, 该数目为 n^N, 所以所述的不等式获证. (关于采用 "两种不同算法" 的大量例子和习题, 以及对组合学方法的众多成果的介绍, 例如, 可见 [61],[45],[103],[104], [122] 等参考文献.)

§2. 概率结构与概念

一切概率统计研究的基础是概率空间或概率模型 (第二章第 1 节)

$$(\Omega, \mathscr{F}, \mathbf{P})$$

其中

Ω 是所有试验结果 ω 的空间,

$\mathscr{F}$ 是子集合的 σ-代数,

$\mathbf{P}$ 是 $\mathscr{F}$ 中的集合的概率测度, 亦即集合 $A \in \mathscr{F}$ 的 σ-可加函数, 有 $0 \leqslant \mathbf{P}(A) \leqslant 1$, $\mathbf{P}(\varnothing) = 0$, $\mathbf{P}(\Omega) = 1$.

- 除了 σ-代数作为概率空间定义中的本质要素被引入概率论之外, 在概率论中还会遇到其他各种不同的子集类: 代数, 可分 σ-代数, 单调类, π-系, λ-系, π-λ-系, 柱集类等等. 参阅第二章第 2 节.

- 事件 (集合)A 与 B 称为相互独立的, 如果 $\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$.

  $\mathscr{F}$ 中的集合所构成的两个子集类 $\mathscr{G}_1$ 与 $\mathscr{G}_2$ 称为相互独立的, 如果任何集合 $A \in \mathscr{G}_1$ 与 $B \in \mathscr{G}_2$ 都是相互独立的.

事件 $A_1, \cdots, A_n$ 称为相互独立, 如果对任何 $k = 1, \cdots, n$ 与 $1 \leqslant i_1 < \cdots < i_k \leqslant n$, 都有

$$\mathbf{P}(A_{i_1} \cap \cdots \cap A_{i_k}) = \mathbf{P}(A_{i_1}) \cdots \mathbf{P}(A_{i_k}).$$

可以相应地定义 $\mathscr{F}$ 中的子集类 $\mathscr{G}_1, \cdots, \mathscr{G}_n$ 的相互独立性.

- 以 $(E, \mathscr{E})$ 表示可测空间, 即赋了 σ-代数 $\mathscr{E}$ 的集合 E.

  重要可测空间的例子有 (参阅第二章第 2 节):

  $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$: 赋了博雷尔 σ-代数 $\mathscr{B}(\mathbb{R})$ 的实直线 $\mathbb{R}$;

  $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n))$: 赋了 σ-代数 $\mathscr{B}(\mathbb{R}^n) = \mathscr{B}(\mathbb{R}) \otimes \cdots \otimes \mathscr{B}(\mathbb{R})$ 的空间 $\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R}$;

  $(\mathbb{R}^\infty, \mathscr{B}(\mathbb{R}^\infty))$: 赋了由柱集类生成的 σ-代数 $\mathscr{B}(\mathbb{R}^\infty)$ 的空间 $\mathbb{R}^\infty = \mathbb{R} \times \cdots \times \mathbb{R} \times \cdots$;

  $(\mathbb{R}^T, \mathscr{B}(\mathbb{R}^T))$: 其中 T 为任一集合, 由所有定义在 T 上的实值函数所构成的空间 $\mathbb{R}^T$, 并赋了由柱集类生成的 σ-代数 $\mathscr{B}(\mathbb{R}^T)$.

  $(\mathcal{C}, \mathscr{B}(\mathcal{C}))$: 由 (定义在区间 $[0,1]$ 或区间 $[0,\infty)$ 上的) 连续函数构成的空间 $\mathcal{C}$, 赋了由开集类所生成的 σ-代数 $\mathscr{B}(\mathcal{C})$, 此处亦即为柱集类所生成的 σ-代数;

  $(\mathcal{D}, \mathscr{B}(\mathcal{D}))$: 由 (例如, 定义在区间 $[0,1]$ 上的) 左极右连函数, 即 (对一切 $t < 1$) 右连续, (对一切 $t > 0$) 左极限存在的函数所构成的空间 $\mathcal{D}$, 赋了由 (按斯科罗霍德距离的) 开集类所生成的 σ-代数 $\mathscr{B}(\mathcal{D})$, 此处亦恰好即为柱集类所生成的 σ-代数.

- *随机变量* $X = X(\omega)$ 即为在 $(\Omega, \mathscr{F})$ 上给出的, 取值于 $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ 的实值函数, 具有 $\mathscr{F}$ 可测性, 即对任何博雷尔集合 $B \in \mathscr{B}(\mathbb{R})$, 都有

$$\{\omega: \ X(\omega) \in B\} \in \mathscr{F}.$$

随机变量的最简单而重要的例子是集合的示性函数: $X(\omega) = I_A(\omega)$, 其中 $A \in \mathscr{F}$, 而

$$I_A(\omega) = \begin{cases} 1, & \text{如果 } \omega \in A, \\ 0, & \text{如果 } \omega \notin A. \end{cases}$$

随机元. 设 $(\Omega, \mathscr{F})$ 与 $(E, \mathscr{E})$ 是两个可测空间, $X = X(\omega)$ 是 Ω 上的取值于 E 的函数. 称 $X(\omega)$ 是随机元, 如果函数 $X(\omega)$ 是 $\mathscr{F}/\mathscr{E}$ 可测的, 意即对任何 $B \in \mathscr{E}$, 都有集合 $\{\omega: \ X(\omega) \in B\} \in \mathscr{F}$.

n 维随机向量 $(X_1(\omega), \cdots, X_n(\omega))$ 是随机变量 $X_1(\omega), \cdots, X_n(\omega)$ 的有序组.

随机序列, 或称*离散时间随机过程* $X = (X_n(\omega))_{n \geqslant 1}$ 就是随机变量的序列 $X_1(\omega), \cdots, X_n(\omega), \cdots$.

随机过程 $X = (X_t(\omega))_{t \in T}$ (具有时间集合 $T \subseteq \mathbb{R}$) 就是随机变量组 $X_t(\omega)$, $t \in T$.

- $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ 上的分布函数$F = F(x)$: 所有 $\mathscr{B}(\mathbb{R})$ 可测的且具有如下三条性质的函数:

  1) $F(x)$ 为非降函数.

  2) $F(-\infty) = 0,\ F(\infty) = 1.$

  3) $F(x)$ 右连续, 且在每个点 $x \in \mathbb{R}$ 处存在左极限.

  若 $X = X(\omega)$ 是定义在概率空间 $(\Omega, \mathscr{F}, \mathbf{P})$ 上的随机变量, 则在 $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ 上按下式定义了一个概率测度 P_X:

$$P_X(B) = \mathbf{P}\{\omega:\ X(\omega) \in B\}$$

(更为准确, 但有些累赘的写法是 $\mathbf{P}(\{\omega:\ X(\omega) \in B\})$), 该测度称为随机变量 $X = X(\omega)$ 的*概率分布*. 函数 $F_X(x) = P_X((-\infty, x])$, 是 $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ 上的分布函数, 称为随机变量 $X = X(\omega)$ 的*分布函数*.

如果 $X = (X_t(\omega))_{t\in T}$ 是随机过程, 则概率 $(t_1 < \cdots < t_n,\ t_i \in T)$

$$P_{t_1,\cdots,t_n}(B) = \mathbf{P}\{\omega:\ (X_{t_1}(\omega), \cdots, X_{t_n}(\omega)) \in B\}, \quad B \in \mathscr{B}(T)$$

称为*随机过程 X 的有限维概率分布*. 函数

$$F_{t_1,\cdots,t_n}(x_1, \cdots, x_n) = \mathbf{P}\{\omega:\ X_{t_1}(\omega) \leqslant x_1, \cdots, X_{t_n}(\omega) \leqslant x_n\}$$

称为*随机过程 X 的有限维分布函数*.

- 如果以勒贝格测度 $\lambda = \lambda(dx)$ 作为 $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ 上的基础测度, 那么 "勒贝格分解"(参阅第三章第 9 节 (29) 式或第七章第 6 节 (3) 式) 将导致如下结论: $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ 上的任一分布函数 $F = F(x)$ 都具有表达式

$$F(x) = aF_{\mathrm{abc}}(x) + F_{\mathrm{sing}}(x),$$

其中 $a \geqslant 0,\ b \geqslant 0,\ a + b = 1,$

$F_{\mathrm{abc}}(x) = \int\limits_{-\infty}^{x} f(y)\lambda(dy)$ 为绝对连续的分布函数 (依博雷尔可测), 具有密度函数 $f = f(y)$ $(f(y) \geqslant 0,\ \int\limits_{-\infty}^{\infty} f(y)\lambda(dy) = 1)$;

$F_{\mathrm{sing}}(x)$ 是奇异的分布函数 (在 $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ 上与其相应的测度 P_{sing}) 是关于勒贝格测度 λ 奇异的测度, 即 $P_{\mathrm{sing}} \perp \lambda$.

奇异分布函数 $F_{\mathrm{sing}}(x)$ 本身又可以进一步表示为

$$F_{\mathrm{sing}}(x) = d \cdot F_{\text{d-sing}}(x) + c \cdot F_{\text{c-sing}}(x),$$

其中 $d \geqslant 0,\ c \geqslant 0,\ d + c = 1,\ F_{\text{d-sing}}(x)$ 为离散的分布函数, 其相应测度 $P_{\text{d-sing}}$ 的支撑集中在不多于可数个点上, 而 $F_{\text{c-sing}}(x)$ 是连续的分布函数, 其相应测度

$P_{\text{c-sing}}$ 的支撑是不可数的勒贝格测度 λ 为零的集合. (函数 $F_{\text{c-sing}}(x)$ 的一个例子是康托尔函数; 参阅第二章第 3 节第 1 小节).

我们记得, $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ 上的测度 μ 的支撑的定义是:

$$\mathrm{supp}(\mu) = \{x \in \mathbb{R}:\ \mu\{y:\ |y - x| \leqslant r\} > 0,\ \forall r > 0\}.$$

由上述讨论可以得出 $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ 上的一切分布函数 $F = F(x)$ 所共同具有的典则表达式 (参阅第二章第 3 节第 18 题):

$$F(x) = \alpha_1 F_{\mathrm{d}}(x) + \alpha_2 F_{\mathrm{abc}}(x) + \alpha_3 F_{\mathrm{sc}}(x),$$

其中, F_{d} $(= F_{\text{d-sing}})$ 为离散的分布函数, F_{abc} 为绝对连续的分布函数, 而 F_{sc} $(= F_{\text{c-sing}})$ 为连续的奇异分布函数, 且 $\alpha_i \geqslant 0,\ i = 1, 2, 3,\ \alpha_1 + \alpha_2 + \alpha_3 = 1.$

§3. 概率论的解析工具与方法

- 定义在概率空间 $(\Omega, \mathscr{F}, \mathbf{P})$ 上的随机变量 $X = X(\omega)$ 的一个重要数字特征是它的数学期望 $\mathbf{E}X$.

  若 $X = X(\omega)$ 为非负随机变量, 则数学期望 $\mathbf{E}X$ 就定义为 $X(\omega)$ 的关于测度 $\mathbf{P}$ 的勒贝格积分:

$$\mathbf{E}X = \int_{\Omega} X(\omega)\mathbf{P}(d\omega).$$

  若 $X = X(\omega)$ 为任意随机变量 $(X = X^+ - X^-$, 其中 $X^+ = \max(X, 0),$ $X^- = -\min(X, 0))$, 则当 $\mathbf{E}X^+$ 与 $\mathbf{E}X^-$ 之一有限 (即 $\min(\mathbf{E}X^+, \mathbf{E}X^-) < \infty$) 时, 我们说数学期望 $\mathbf{E}X$*存在*, 或说*确定*. 此时, 根据定义, 我们令

$$\mathbf{E}X = \mathbf{E}X^+ - \mathbf{E}X^-.$$

  若 $\mathbf{E}X^+ < \infty$ 且 $\mathbf{E}X^- < \infty$, 亦即 $\mathbf{E}|X| < \infty$, 因为 $|X| = X^+ + X^-$, 则说数学期望 $\mathbf{E}X$*有限*(或说随机变量 X*可积*). (参阅第二章第 6 节).

- 概率论中的重要解析工具有: 关于积分号下取极限的各种定理 (单调收敛定理, 法图引理, 勒贝格控制收敛定理), 一致可积性概念, 不等式 (切比雪夫不等式, 柯西–布尼亚科夫斯基不等式, 延森不等式, 李雅普诺夫不等式, 赫尔德不等式, 闵可夫斯基不等式, 等等), 拉东–尼科迪姆定理, 富比尼定理, 勒贝格积分中的换元定理. (参阅第二章第 6 节.)

- 随机变量 $X = X(\omega)$ 的*方差*是

$$\mathbf{D}X = \mathbf{E}(X - \mathbf{E}X)^2.$$

  量 $\sigma = +\sqrt{\mathbf{D}X}$ 称为随机变量 X(偏离 $\mathbf{E}X$) 的*标准 (线性) 偏差*.

如果 (X, Y) 为随机变量对, 则它们的*协方差*是:

$$\mathbf{cov}(X, Y) = \mathbf{E}(X - \mathbf{E}X)(Y - \mathbf{E}Y)$$

(假定相应的数学期望是确定的).

如果 $0 < \mathbf{D}X < \infty$, $0 < \mathbf{D}Y < \infty$, 则量

$$\rho(X, Y) = \frac{\mathbf{cov}(X, Y)}{\sqrt{\mathbf{D}X}\sqrt{\mathbf{D}Y}}$$

称为随机变量 X 与 Y 的*相关系数*.

若 X 为随机变量, 则数学期望 $\mathbf{E}X^n$(若其确定) 称为 X 的 n 阶矩或第 n 个矩. 量 $\mathbf{E}(X)_n = \mathbf{E}X(X-1)\cdots(X-n+1)$ 称为其 n 阶阶乘矩.

- 若 $F = F(x)$ 为分布函数, 则函数

$$\varphi(t) = \int_{\mathbb{R}} e^{itx} dF(x) \quad \left(= \int_{\mathbb{R}} \cos tx dF(x) + i \int_{\mathbb{R}} \sin tx dF(x)\right), \quad t \in \mathbb{R},$$

称为分布 F 的*特征函数*.

当 $F = F_X$ 是随机变量 X 的分布函数时, 函数

$$\varphi_X(t) = \int_{\mathbb{R}} e^{itx} dF_X(x), \quad t \in \mathbb{R},$$

称为*随机变量* $X = X(\omega)$ *的特征函数*.(参阅第二章第 12 节.) 此时 $\varphi_X(t) = \mathbf{E}e^{itX(\omega)}$.

- 若 X 是非负随机变量, 具有分布函数 $F_X = F_X(x)$, 则函数

$$\widehat{F}_X(\lambda) = \int_0^\infty e^{-\lambda x} dF_X(x), \quad \lambda > 0,$$

称为分布函数 F_X 或随机变量 X 的*拉普拉斯变换*. 在此, 有 $\widehat{F}_X(\lambda) = \mathbf{E}e^{-\lambda X}$.

在第二章第 3 节中列举了最常用的离散分布或具有密度的概率分布.

- 研究离散随机变量概率性质的最有效的工具是*母函数*. 这种方法在其他数学分支中也广为所知, 并且是研究具有复杂结构的数列性质的有力工具.

如果随机变量 X 分别以概率 $p_0, p_1, p_2, \cdots$ $(p_k \geqslant 0, \sum_{k=0}^{\infty} p_k = 1)$ 取值 $0, 1, 2, \cdots$, 那么在概率论中, 就将 X 的母函数 $G(s)$ 定义为

$$G(s) = \mathbf{E}s^X \quad \left(= \sum_{k=0}^{\infty} p_k s^k\right), \quad |s| \leqslant 1.$$

由随机变量 X 的母函数 $G(s)$, 可以唯一确定其分布 $(p_k)_{k\geqslant 0}$:

$$p_k = \mathbf{P}\{X = k\} = \frac{G^{(k)}(0)}{k!}.$$

若有向量值随机变量 $X = (X_1, \cdots, X_d)$, 其每个分量 X_i 均在集合 $\{0, 1, 2, \cdots\}$ 中取值, 则其母函数 $G(s)$, 其中 $s = (s_1, \cdots, s_d)$, 按如下方式定义:

$$G(s_1, \cdots, s_d) = \mathbf{E}s_1^{X_1} \cdots s_d^{X_d} = \sum_{k_1, \cdots, k_d = 0}^{\infty} p_{k_1, \cdots, k_d} s_1^{k_1} \cdots s_d^{k_d},$$

其中 $p_{k_1, \cdots, k_d} = \mathbf{P}\{X_1 = k_1, \cdots, X_d = k_d\}$, $|s_j| \leqslant 1$, $j = 1, \cdots, d$.

如果随机变量 $X_1, \cdots, X_d$ 相互独立, 则有

$$G(s_1, \cdots, s_d) = G_1(s_1) \cdots G_d(s_d),$$

其中 $G_j(s_j) = \mathbf{E}s_j^{X_j}$, $j = 1, \cdots, d$.

上面所给出的关于随机变量 X 的母函数 $G(s)$ 的定义中, 假定了 X 只在集合 $\{0, 1, 2, \cdots\}$ 中取值, 其值非负. 为了满足不同的需求, 可将这一概念推广到 X 既可取非负整数值, 又可取负整数值的情形, 亦即有 $\mathbf{P}\{X = k\} = p_k$, 其中 $k = 0, \pm 1, \pm 2, \cdots$, 且 $\sum_{k=-\infty}^{\infty} p_k = 1$.

对于这样的随机变量 X, 其母函数 $G(s)$ 的定义为 (类似的情形 $k = 0, 1, \cdots$)

$$G(s) = \mathbf{E}s^X = \sum_{k=-\infty}^{\infty} p_k s^k$$

(其中 s 使得 $\mathbf{E}|s^X| < \infty$).

此类随机变量的典型例子是: $X = X_1 - X_2$, 其中随机变量 X_1 与 X_2 均在集合 $\{0, 1, 2, \cdots\}$ 中取值, 分别有自己的母函数 $G_{X_1}(s)$ 与 $G_{X_2}(s)$.

若 X_1 与 X_2 相互独立, 则有

$$G_X(s) = G_{X_1}(s) G_{X_1}\left(\frac{1}{s}\right).$$

特别地, 若 X_i 服从参数为 λ_i 的泊松分布, $i = 1, 2$, 则 $G_{X_i}(s) = e^{-\lambda_i(1-s)}$, 因而随机变量 $X = X_1 - X_2$ 的母函数 $G_X(s)$ 就是

$$G_X(s) = e^{-(\lambda_1+\lambda_2)+\lambda_1 s+\lambda_2/s} = e^{-(\lambda_1+\lambda_2)} e^{\sqrt{\lambda_1\lambda_2}(t+1/t)},$$

其中 $t = s\sqrt{\lambda_1/\lambda_2}$.

由分析数学知道, 对 $x \in \mathbb{R}$, 有

$$e^{x(t+1/t)} = \sum_{k=-\infty}^{\infty} t^k I_k(2x),$$

其中 $I_k(2x)$ 是 k 阶第一类修正的贝塞尔函数 (例如, 参阅 [79] 第 5 卷, p820 ~ 825):

$$I_k(2x) = x^k \sum_{r=0}^{\infty} \frac{x^{2r}}{r!\,\Gamma(k+r+1)}, \quad k = 0, \pm 1, \pm 2, \cdots$$

(换句话说, 对于每个 $x \in \mathbb{R}$, 序列 $\{I_k(2x)\}_{k=0,\pm1,\cdots}$ 的母函数由 $e^{x(t+1/t)}$ 的公式给出.)

这样一来, 随机变量 $X = X_1 - X_2$ 的概率分布就由如下的公式给出:

$$\mathbf{P}\{X = k\} = e^{-(\lambda_1+\lambda_2)}\left(\frac{\lambda_1}{\lambda_2}\right)^{k/2} I_n(2\sqrt{\lambda_1\lambda_2}),$$

其中 $k = 0, \pm1, \pm2, \cdots$.

关于与母函数有关的其他的计算的例子可以参阅, 例如, 第二章第 6 节第 28 题, 第 32 题, 第 8 节第 22 题, 第七章第 2 节第 18 题, 第八章第 8 节第 19 题.

- 正如前面所指出的, 母函数在数学的许多分支中都起着重要的作用, 特别地, 在离散数学与组合论中亦是如此.

从本质上说, 解决组合问题的代数方法的基础就是母函数, 而这恰恰形成了组合学中的一个方向, 即所谓的*代数组合学*.

简单说来, 事情是这样的, 许多组合运算和组合解释都能够与一定的代数运算和代数解释挂上钩.

我们以彩票为例, 看看如何可以利用母函数的代数性质来解决组合计数问题. 假定这些彩票具有自 000000 至 999999 这样一些 6 位数号码. 要求求出被人随机购买的一张彩票的号码的前 3 位数字之和与后 3 位数字之和相等的概率. 显然, 这里对有利样本点个数的计算是一个纯粹的组合计数问题. 例如, 可以通过枚举的办法来计数. 然而, (正如我们后面将会看见的,) 该数目为 55252[①], 对于手算, 而不是用计算机来算, 其难度可想而知.

在这个涉及 6 位数票号的问题中, 如同在类似的 $2n$ 个数码问题中一样, 用母函数办法可以很快地求得答案. 设 $X = (X_1, \cdots, X_6)$ 为 6 维随机向量, 其各个分量相互独立, 均有 $p_k = \mathbf{P}\{X_i = k\} = 1/10$, 其中 $k = 0, 1, \cdots, 9$.

X_i 的母函数为

$$G_{X_i}(s) = \sum_{k=0}^{9} p_k s^k = \frac{1}{10}(1 + s + \cdots + s^9) = \frac{1 - s^{10}}{10(1-s)}.$$

由独立性可得

$$G_{X_1+X_2+X_3}(s) = G_{X_1}(s)G_{X_2}(s)G_{X_3}(s) = \left(\frac{1-s^{10}}{10(1-s)}\right)^3.$$

而 $G_{X_4+X_5+X_6}(s)$ 亦同样如此.

现在考察随机变量 $Y = (X_1 + X_2 + X_3) - (X_4 + X_5 + X_6)$.

由独立性立知

$$G_Y(s) = G_{X_1+X_2+X_3}(s)G_{X_4+X_5+X_6}\left(\frac{1}{s}\right) = \frac{1}{10^6}\frac{1}{s^{27}}\left(\frac{1-s^{10}}{1-s}\right)^6.$$

①原书此处写为 55250, 与下面的计算结果不符 —— 译者注.

如所周知, 在母函数

$$G_Y(s) = \sum_k q_k s^k$$

中, (s^0 的) 系数 q_0 就是概率 $\mathbf{P}\{Y = 0\}$, 而它正是我们所感兴趣的 " 彩票号码的前 3 位数字之和与后 3 位数字之和相等" 的概率.

分别将 $(1-s^{10})^6$, $(1-s)^{-6}$ 和 $\frac{1}{s^{27}}\left(\frac{1-s^{10}}{1-s}\right)^6$ 展为 (关于 s^k, $k = 0, \pm1, \pm2, \cdots$ 的) 幂级数, (经过简单的但却足够烦杂的 "组合数" 运算,) 即可求得

$$q_0 = \frac{55252}{10^6} = 0.055252.$$

(亦可参阅第二章第 6 节第 79 题.)

通常情形下, 任意一个数列 $a = (a_n)_{n\geqslant0}$ 的*母函数*是指形如下式的 (形式) 幂级数

$$G_a(x) = a_0 + a_1x + a_2x^2 + \cdots, \quad x \in \mathbb{R}.$$

若该级数有非 0 的收敛半径, 则它事实上定义了一个*函数*, 它的性质可以反映出数列 $a = (a_n)_{n\geqslant0}$ 的许多性质. 关于此类级数的收敛半径问题通常会引起关注, 然而在关于母函数的广义理论中, 却认为形式级数 $G_a(x)$ 是数列 $a = (a_n)_{n\geqslant0}$ 的一种通过双方对应

$$(a_n) \quad \leftrightarrow \quad G_a(x)$$

建立起来的形式独特的 "重新编码". 根据这种理解, 那么如果 $(b_n) \leftrightarrow G_b(x)$, 而 c 为常数, 则有

$$(a_n + cb_n) \quad \leftrightarrow \quad G_a(x) + cG_b(x).$$

对应 " $\leftrightarrow$ " 的最重要的性质之一是

$$\left(\sum_{i=0}^{n} a_i b_{n-i}\right)_{n\geqslant0} \quad \leftrightarrow \quad G_a(x)G_b(x),$$

这表明, 数列 $a = (a_n)_{n\geqslant0}$ 与 $b = (b_n)_{n\geqslant0}$ 的卷积对应着母函数的乘积. 不难看出, 所引入的形式运算 (相加, 乘以常数, 形式级数的相乘) 都具有交换律, 结合律和分配律, 一言以蔽之, 形式级数具有*代数结构*. (更详细的, 可参阅 [61], [122], [103], [104].)

与根据数列 $a = (a_n)_{n\geqslant0}$ 构造幂级数 $G_a(x)$ 相对应, 有时建立所谓的*指数母函数*则更为有益:

$$E_a(x) = \sum_{n\geqslant0} a_n \frac{x^n}{n!}.$$

(此处仍将其理解为形式幂级数.) 如同母函数一样, 亦存在如下的一一对应关系:

$$(a_n) \quad \leftrightarrow \quad E_a(x);$$

并具有性质:

$$\begin{aligned}(a_n+cb_n) &\quad\leftrightarrow\quad E_a(x)(cE_b(x)),\\ \left(\sum_{i=0}^{n}C_n^i a_i b_{n-i}\right)_{n\geqslant 0} &\quad\leftrightarrow\quad E_a(x)E_b(x).\end{aligned}$$

看几个例子. 如果数列 $a=(a_n)_{n\geqslant 0}$ 为 $a_n=1,\ n\geqslant 0$, 则其母函数形如

$$G_a(x)=\sum_{n=0}^{\infty}x^n\quad\left(=\frac{1}{1-x},\quad |x|<1\right),$$

且有

$$[G_a(x)]^N=\left[\sum_{n=0}^{\infty}X^n\right]^N=\sum_{n=0}^{\infty}C_{N+n-1}^n x^n,$$

该式可按下法得到.

观察 $(1+x+x^2+\cdots)^N$. 我们可有几种办法得到 x^n 的系数? 写

$$(1+x+x^2+\cdots)^N=(1+x+x^2+\cdots)(1+x+x^2+\cdots)\cdots(1+x+x^2+\cdots).$$

如果从右端的第一个括弧中取出 x^{n_1}, 从第二个括弧中取出 $x^{n_2},\cdots$, 从第 N 个括弧中取出 x^{n_N}, 则应当有 $x^{n_1}x^{n_2}\cdots x^{n_N}=x^n$. 满足关系式 $n_1+n_2+\cdots+n_N=n$ 的非负整数组 $(n_1,n_2,\cdots,n_N)$ 的数目等于该方程的非负整数解的数目, 即为 C_{n+N-1}^n(参阅第一章第 1 节第 6 题). 如此一来, 即知数列 $(C_{N+n-1}^n)_{n\geqslant 0}$ 的母函数是 $(1-x)^{-N},\ |x|<1$.

由此可知

$$\sum_{n=0}^{\infty}2^{-n}C_{N+n-1}^n=2^N.$$

数列 $(C_N^n)_{n\geqslant 0}$(当 $n>N$ 时, $C_N^n=0$) 的母函数是 $(1+x)^N$:

$$(1+x)^N=\sum_{n=0}^{N}C_N^n x^n.$$

(可采用上面证明 $(C_{N+n-1}^n)_{n\geqslant 0}$ 的母函数的类似办法证明这一结论.)

考察表达式

$$(1+x)^N(1+x)^M=(1+x)^{N+M},$$

将其两端按 x 的方幂展开, 再令两端 x^k 的系数相等, 容易得到

$$\sum_{j=0}^{N}C_N^j C_M^{k-j}=C_{N+M}^k.$$

(该等式以 “范德蒙德卷积” 的名字著称, 或称为超几何等式; 亦可参阅第一章第 2 节第 2 题.) 所采用的证明方法很好地展示了母函数方法可以用来建立各式各样的组合关系式.

在关于 “母函数” 这部分内容的末尾, 我们回过来观察前面提到过的二阶斯特林数 S_N^n 与贝尔数 B_N. 我们记得, 二阶斯特林数 S_N^n 是将 N 元集合 $\boldsymbol{A}$ 划分为 n 个两两不交的非空集合的所有可能的分割 $\mathscr{D}$ 的数目. 贝尔数 $B_N=\sum\limits_{n=1}^{N}S_N^n$ 是对 N 元集合 $\boldsymbol{A}$ 所可能作的所有不同的分割的数目.

在第 1 节 “组合论基础” 中, 曾引入公式 $n^N=\sum\limits_{k=1}^{n}S_N^k(n)_k$, 并给出了它的组合学证明. (我们记得, $S_N^1=S_N^N=1,\ S_N^0=0$, 且对 $n>N$, 有 $S_N^n=0$.)

由该公式可以推出, 对于每个 $N\geqslant 1$, 多项式

$$P_N(x)=x^N-\sum_{n=1}^{N}S_N^n(x)_n,\quad x\in\mathbb{R}$$

均以 $x=1,\cdots,N$ 为根. 又由于 $x=0$ 也是它的根, 故知 $P_N(x)\equiv 0$. 如此一来, 对任何 $N\geqslant 1$ 与 $x\in\mathbb{R}$, 就都有

$$x^N=\sum_{n=1}^{N}S_N^n(x)_n.$$

若再认为对一切 $N\geqslant 1$, 都有 $S_N^0=0$, 并假定 $S_0^0=1$ 与 $(x)_0=1$, 则可见

$$x^N=\sum_{n=0}^{N}S_N^n(x)_n$$

对一切 $N=0,1,2,\cdots$ 与 $x\in\mathbb{R}$ 成立.

利用这一公式, 我们得到

$$\begin{aligned}\sum_{n\geqslant 0}\left(\sum_{N\geqslant 0}S_N^n\frac{y^N}{N!}\right)(x)_n&=\sum_{N\geqslant 0}\frac{y^N}{N!}\left(\sum_{n\geqslant 0}S_N^n(x)_n\right)\\&=\sum_{N\geqslant 0}\frac{(yx)^N}{N!}=e^{yx}=(1+(e^y-1))^x=\sum_{n\geqslant 0}\frac{1}{n!}(e^y-1)^n(x)_n,\end{aligned}$$

其中最后一步用到泰勒公式 $(1+z)^x=\sum_{n\geqslant 0}\frac{z^n}{n!}(x)_n$.

比较上式的最左端与最右端, 我们得到二阶斯特林数 $(S_N^n)_{N\geqslant 0}$ 的指数型母函数: 对一切 $n\geqslant 0$, 有

$$\sum_{N\geqslant 0}S_N^n\frac{y^N}{N!}=\frac{1}{n!}(e^y-1)^n$$

$(S_0^0=1$; 对 $N\geqslant 1$, 有 $S_N^0=0$; 对 $N<n$, 有 $S_N^n=0)$.

用类似的方法, 我们还可以得到数列 $(S_N^n)_{n\geqslant 0}$ 的母函数:

$$\sum_{n\geqslant 0} S_N^n x^n = e^{-x}\sum_{m\geqslant 0}\frac{m^N x^m}{m!}.$$

事实上,

$$x^n e^x = \sum_{i\geqslant 0}\frac{x^{i+n}}{i!} = \sum_{m\geqslant 0}(m)_n\frac{x^m}{m!},$$

此因当 $m\leqslant n-1$ 时, 有 $(m)_n=0$. 再利用等式 $m^N=\sum\limits_{n\geqslant 0} S_N^n(m)_n$, 得

$$e^x\sum_{n\geqslant 0} S_N^n x^n = \sum_{n\geqslant 0} S_N^n x^n e^x = \sum_{n\geqslant 0} S_N^n\left(\sum_{m\geqslant 0}(m)_n\frac{x^m}{m!}\right)$$
$$= \sum_{m\geqslant 0}\frac{x^m}{m!}\left(\sum_{n\geqslant 0} S_N^n(m)_n\right) = \sum_{m\geqslant 0}\frac{m^N x^m}{m!},$$

这就是所要证明的.

若令 $x=1$, 则可得到关于贝尔数的*多宾斯基公式*:

$$B_N = e^{-1}\sum_{m\geqslant 0}\frac{m^N}{m!}.$$

在前面定义二阶斯特林数 S_N^n 时, 我们是从其组合意义出发的, 即把它视为将 N 元集合 $\boldsymbol{A}$ 划分为 n 个两两不交的非空集合的所有可能的分割 $\mathscr{D}$ 的数目. 然后再证明对一切 $N\geqslant 0$, 都有 $x^N=\sum\limits_{n=0}^{N} S_N^n(x)_n$.

一阶斯特林数 $(s_N^n)_{0\leqslant n\leqslant N}$ 可以按照下式代数地定义:

$$(x)_N = \sum_{n=0}^{N} s_N^n x^n. \qquad (*)$$

这些数的组合意义如下: 设 $(\pi_1,\cdots,\pi_N)$ 是正整数 $(1,\cdots,N)$ 的任意一个排列. 以 c_N^n 表示其中有 n 个 “轮换” 的排列个数. (如, 在排列 $\begin{pmatrix}1, & 2, & 3, & 4, & 5\\ 2, & 1, & 4, & 5, & 3\end{pmatrix}$ 中有两个 “轮换”.) 可以证明如下的递推关系式:

$$c_N^n = c_{N-1}^{n-1} + (N-1)c_{N-1}^n$$

(其中 $c_0^0=1$), 从中可以推出

$$\sum_{n=0}^{N} c_N^n x^n = x(x+1)\cdots(x+N-1).$$

将这个关于数 $c_N^0, c_N^1, \cdots, c_N^N$ 的母函数与上面的关于一阶斯特林数 $s_N^0, s_N^1, \cdots, s_N^N$ 的母函数 $(*)$ 相对照, 可知

$$c_N^n = (-1)^{N-n} s_N^n.$$

如此一来, 除了符号之外, 一阶斯特林数就重合于正整数 $(1,\cdots,N)$ 的恰含有 n 个 “轮换” 的排列的个数.

若 X 为在集合 $N=\{0,1,2,\cdots\}$ 中取值的离散随机变量, 则其母函数曾定义为 $G(s)=\mathbf{E}s^X$, $|s|\leqslant 1$. 现在记 $p_k=\mathbf{P}\{X=k\}$, 我们发现, 幂级数

$$G(s) = \sum_{k=0}^{\infty} p_k s^k$$

就是数列 $(p_k)_{k\geqslant 0}$ 的母函数.

与此相关联的概念是*矩母函数*(参阅第二章第 6 节第 32 题)

$$M(s) = \mathbf{E}e^{sX}$$

(数学期望 $\mathbf{E}e^{sX}$ 显然是确定的与有限的, 如果, 例如, $-1<s<0$). 当所有的矩 $m^{(k)}=\mathbf{E}X^k$, $k\geqslant 1$, 都有限时, 由 $M(s)$ 的定义可知, (形式) 级数

$$M(s) = \sum_{k=0}^{\infty} m^{(k)}\frac{s^k}{k!}$$

是数列 $(m^{(k)})_{k\geqslant 0}$ 的指数型母函数. 概率论中, 在考察矩 $m^{(k)}=\mathbf{E}X^k$ 的同时, 也往往考察*阶乘矩*

$$(m)_k = \mathbf{E}(X)_k \equiv \mathbf{E}X(X-1)\cdots(X-k+1)$$

与*二项矩*

$$b_{(k)} = \mathbf{E}\frac{(X)_k}{k!} = \frac{(m)_k}{k!}$$

(二项矩名称的来源是: 由于 $C_X^k=\frac{(X)_k}{k!}$, 所以 $b_{(k)}=\mathbf{E}C_X^k$ 是二项式系数的期望).

与阶乘矩数列 $((m)_k)_{k\geqslant 0}$ 相对应的指数型母函数为

$$(M)(s) = \sum_{k=0}^{\infty}(m)_k\frac{s^k}{k!},$$

与 $(b_{(k)})_{k\geqslant 0}$ 相对应的母函数为

$$B(s) = \sum_{k=0}^{\infty} b_{(k)} s^k.$$

显然有①

$$M(s)=G(e^s),\quad (M)(s)=B(s)=M(s+1).$$

了解下述事实是有益的: 利用前面介绍过的公式

$$x^N=\sum_{n=0}^{N}S_N^n(x)_n,\quad (x)_N=\sum_{n=0}^{N}s_N^n x^n,$$

其中 S_N^n 与 s_N^n 分别为二阶斯特林数与一阶斯特林数, 可以推出联系矩 $m^{(n)}=\mathbf{E}X^n$ 与阶乘矩 $(m)_n=\mathbf{E}(X)_n$, $n\geqslant 0$ 的关系式:

$$m^{(N)}=\sum_{n=0}^{N}S_N^n(m)_n,\quad (m)_N=\sum_{n=0}^{N}s_N^n m^{(n)}.$$

• 应当指出, 数学分析中广为了解的许多数 (诸如, 伯努利数、欧拉数, 等等) 与许多多项式 (伯努利多项式、欧拉多项式、埃尔米特多项式、阿贝尔多项式, 等等) 都是通过母函数来定义的.

(a) 伯努利数 $b_0,b_1,b_2,\cdots$ 与伯努利多项式 $B_0(x),B_1(x),B_2(x),\cdots$ 是通过 (它们自己的) 指数型母函数来定义的:

$$\frac{s}{e^s-1}=\sum_{n=0}^{\infty}b_n\frac{s^n}{n!}\quad 与\quad \frac{se^{sx}}{e^s-1}=\sum_{n=0}^{\infty}B_n(x)\frac{s^n}{n!}.$$

(一些特殊的值为: $b_0=1,b_1=-\frac{1}{2},b_2=\frac{1}{6},b_4=-\frac{1}{30},b_6=\frac{1}{42},b_8=-\frac{1}{30}$ 与 $B_0(x)=1,B_1(x)=x-\frac{1}{2},B_2(x)=x^2-x+\frac{1}{6},B_3(x)=x^3-\frac{3}{2}x^2+\frac{1}{2}x$, 例如, 可参阅 [28]). 除了 $b_1=-\frac{1}{2}$ 外, 其余脚标为奇数的伯努利数都是 0. 关于它们的一些其他性质开列于下:

(i) $b_N=\sum\limits_{n=0}^{N}C_N^n b_{N-n},\ N=2,3,\cdots$.

(ii) 所有 b_N 均为有理数.

(iii) $B_N(0)=b_N,\ B_N(1)=(-1)^N b_N,\ N\geqslant 0$.

(iv) $B_N(x)=\sum\limits_{n=0}^{N}C_N^n b_n x^{N-n},\ N\geqslant 1$.

(v) $B_N'(x)=NB_{N-1}(x),\ N\geqslant 1$.

(b) 欧拉数 $e_0,e_1,e_2,\cdots$ 与欧拉多项式 $E_0(x),E_1(x),E_2(x),\cdots$ 是通过 (它们自己的) 指数型母函数来定义的:

$$\frac{2e^s}{e^{2s}+1}=\sum_{n=0}^{\infty}e_n\frac{s^n}{n!}\quad 与\quad \frac{2e^{sx}}{e^s+1}=\sum_{n=0}^{\infty}E_n(x)\frac{s^n}{n!}.$$

①原书写为: $(M)(s)=B(s)=G(s+1)$ —— 译者注.

由于 $\frac{2e^s}{e^{2s}+1}=\frac{1}{\operatorname{ch}s}$, 所以可以说, 欧拉数 $e_0,e_1,e_2,\cdots$ 的指数型母函数就是函数 $\frac{1}{\operatorname{ch}s}$.

由定义可知:

(i) $e_N=2^N E_N\left(\frac{1}{2}\right),\ N\geqslant 0$.

(ii) $E_N(x)=\sum\limits_{n=0}^{N}C_N^n E_n\dfrac{1}{2^n}\left(x-\dfrac{1}{2}\right)^{N-n},\ N\geqslant 0$.

(iii) $E_N'(x)=NE_{N-1}(x),\ N\geqslant 1$.

(iv) 脚标为奇数的欧拉数都是 0, 脚标为偶数的欧拉数都是整数.

一些特殊的欧拉数为: $e_0=1,e_2=-1,e_4=5,e_6=-61,e_8=1385$; 参阅 [28].

(c) 埃尔米特多项式的引入方式在分析中与在概率论中多少有些不同.

在分析中, 埃尔米特多项式 $H_n(x)$, $n\geqslant 0$, 是按照通常的方式, 由公式定义的:

$$H_n(x)=(-1)^n\frac{D^n\psi(x)}{\psi(x)},$$

其中 $\psi(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2}$. 相应的指数型母函数 ($s\in\mathbb{R}$, $x\in\mathbb{R}$) 具有简洁的表达式:

$$\sum_{n=0}^{\infty}H_n(x)\frac{s^n}{n!}=e^{2sx-s^2}.$$

在概率论中, 一般采用如下的 (多少有些不同的) 埃尔米特多项式

$$\mathrm{He}_n(x)=(-1)^n\frac{D^n\phi(x)}{\phi(x)},\quad n\geqslant 0,$$

其中 $\phi(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$, 即标准正态分布 $\mathcal{N}(0,1)$ 的密度函数. (请注意, 在第二章第 11 节中, 这些多项式 $\mathrm{He}_n(x)$ 写为 $H_n(x)$.)

与多项式 $\mathrm{He}_n(x)$, $n\geqslant 0$, 相应的指数型母函数 ($s\in\mathbb{R}$, $x\in\mathbb{R}$) 具有如下形式:

$$\sum_{n=0}^{\infty}\mathrm{He}_n(x)\frac{s^n}{n!}=e^{2sx-s^2/2}.$$

不难看出

$$\mathrm{He}_n(x)=2^{-n/2}H_n(2^{-1/2}x).$$

一些特殊的值为:

$$\begin{aligned}
&H_0(x)=1, && \mathrm{He}_0(x)=1,\\
&H_1(x)=2x, && \mathrm{He}_1(x)=x,\\
&H_2(x)=4x^2-2, && \mathrm{He}_2(x)=x^2-1,\\
&H_3(x)=8x^3-12x, && \mathrm{He}_3(x)=x^3-3x.
\end{aligned}$$

由于布朗运动 $B=(B_t)_{t\geqslant 0}$, 在随机过程论中经常使用如下形式的埃尔米特多项式 $\mathrm{He}_n(x,t)$, $n\geqslant 0$, $x\in\mathbb{R}$, $t\in\mathbb{R}_+$, 它们是通过指数型母函数 $(s\in\mathbb{R})$ 来定义的:

$$\sum_{n=0}^{\infty}\mathrm{He}_n(x,t)\frac{s^n}{n!}=e^{2sx-\frac{s^2}{2}t}.$$

(往往将 $\mathrm{He}_n(x,t)$ 写为 $H_n(x,t)$). 引入这些对象的原因是: 它们具有某些有趣的性质, 例如, 对于 (标准) 布朗运动 $B=(B_t)_{t\geqslant 0}$, 如下二过程

$$(\mathrm{He}_n(B_t,t))_{t\geqslant 0},\ n\geqslant 0 \quad 与 \quad \left(e^{2sB_t-\frac{s^2}{2}t}\right)_{t\geqslant 0}$$

都是鞅 (关于由布朗运动所生成的滤子).

(d) 设 X 为随机变量, 其矩母函数

$$G(s)=\mathbf{E}e^{sX}$$

对 $|s|<\lambda$ 有限, 其中 λ 为某个正数.

引入函数

$$A(s,x)=\frac{e^{sx}}{G(s)},\quad x\in\mathbb{R},\quad |s|<\lambda.$$

(在保险与金融数学中, 函数 $x\rightsquigarrow\frac{e^{sx}}{G(s)}$ 称为埃绍尔变换, 参阅第七章第 11 节.)

根据函数 $A(s,x)$ 可以通过如下的展开式来定义阿贝尔多项式 (亦称为谢费尔 (Sheffer) 多项式) $Q_0(x),Q_1(x),Q_2(x),\cdots$:

$$A(s,x)=\sum_{k=0}^{\infty}Q_k(x)\frac{s^k}{k!}.$$

换言之,$A(s,x)=\frac{e^{sx}}{\mathbf{E}e^{sX}}$ 就是多项式序列 $(Q_k(x))_{k\geqslant 0}$ 的母函数.

若 X 为服从区间 $[0,1]$ 上均匀分布的随机变量, 则其矩母函数为

$$G(s)=\mathbf{E}e^{sX}=\frac{e^s-1}{s}.$$

此时

$$A(s,x)=\frac{se^{sx}}{e^s-1},$$

因而, 相应的阿贝尔多项式 $Q_k(x)$ 不是别的, 恰恰就是伯努利多项式 $B_k(x)$.

若 X 为伯努利随机变量, 有 $\mathbf{P}\{X=1\}=\mathbf{P}\{X=-1\}=1/2$, 则其矩母函数为

$$G(s)=\mathbf{E}e^{sX}=\frac{e^s+1}{2}.$$

此时

$$A(s,x)=\frac{2e^{sx}}{e^s+1},$$

因而, 相应的阿贝尔多项式就是埃尔米特多项式.

若 X 为服从标准正态分布 $\mathcal{N}(0,1)$ 的随机变量, 则其矩母函数为

$$G(s)=e^{s^2/2}.$$

此时

$$A(s,x)=e^{sx-s^2/2},$$

因而, 相应的阿贝尔多项式 $Q_k(x)$ 重合于埃尔米特多项式 $\mathrm{He}_k(x)$.

以 $\kappa_1,\kappa_2,\cdots$ 表示随机变量 X 的半不变量. 则可证明,

$$\begin{aligned}&Q_0(x)=1,\\&Q_1(x)=x-\kappa_1,\\&Q_2(x)=(x-\kappa_1)^2-\kappa_2,\\&Q_3(x)=(x-\kappa_1)^3-3\kappa_2(x-\kappa_1)-\kappa_3.\end{aligned}$$

当 $X\sim\mathcal{N}(0,1)$ 时, 有 $\kappa_1=0$, $\kappa_2=1$, $\kappa_3=\kappa_4=\cdots=0$, 因而就有

$$\begin{aligned}&Q_0(x)=\mathrm{He}_0(x)=1,\\&Q_1(x)=\mathrm{He}_1(x)=x,\\&Q_2(x)=\mathrm{He}_2(x)=x^2-1,\\&Q_3(x)=\mathrm{He}_3(x)=x^3-3x.\end{aligned}$$

我们指出, 为了能唯一确定多项式 $Q_k(x)$, $k=1,\cdots,n$, 事实上只需要 $\mathbf{E}|X|^n<\infty$. 此时有如下的关系式 (所谓的阿贝尔式) 成立 $(Q_0(x)\equiv 1)$:

$$Q_k'(x)=kQ_{k-1}(x),\quad 1\leqslant k\leqslant n.$$

- 非负随机变量 X 关于子 σ-代数 $\mathscr{G}\subseteq\mathscr{F}$ 的条件数学期望是非负随机变量 (一般来说, 是取值于 $\overline{\mathbb{R}}$ 的广义随机变量) $\mathbf{E}(X\,|\,\mathscr{G})=\mathbf{E}(X\,|\,\mathscr{G})(\omega)$, 有

  1) $\mathbf{E}(X\,|\,\mathscr{G})$ 是 $\mathscr{G}$ 可测的;

  2) 对任何 $A\in\mathscr{G}$, 都有

$$\mathbf{E}[XI_A]=\mathbf{E}[\mathbf{E}(X\,|\,\mathscr{G})I_A].$$

对于任意随机变量 $X\ (=X^+-X^-)$, 若 $\mathbf{P}$-a.s. 地有

$$\min\left[\mathbf{E}(X^+\,|\,\mathscr{G})(\omega),\ \mathbf{E}(X^-\,|\,\mathscr{G})(\omega)\right]<\infty,$$

则认为 X 关于子 σ-代数 $\mathscr{G} \subseteq \mathscr{F}$ 的条件数学期望是确定的, 并且

$$\mathbf{E}(X\,|\,\mathscr{G})(\omega) = \mathbf{E}(X^+\,|\,\mathscr{G})(\omega) - \mathbf{E}(X^-\,|\,\mathscr{G})(\omega).$$

当 $X(\omega) = I_A(\omega)$, 即为某个集合 $A \in \mathscr{F}$ 的示性函数时, 将条件数学期望 $\mathbf{E}(I_A\,|\,\mathscr{G}) = \mathbf{E}(I_A\,|\,\mathscr{G})(\omega)$ 写为 $\mathbf{P}(A\,|\,\mathscr{G})$ 或 $\mathbf{P}(A\,|\,\mathscr{G})(\omega)$, 称其为事件 A 关于子 σ-代数 $\mathscr{G} \subseteq \mathscr{F}$ 的条件概率.

如果 $\mathscr{G}$ 是某个随机元 $Y = Y(\omega)$ 所生成的 σ-代数 (为直观起见, 通常将该 σ-代数记为 $\mathscr{G}_Y$ 或 $\sigma(Y)$), 则相应的 $\mathbf{E}(X\,|\,\mathscr{G}_Y)$ 与 $\mathbf{P}(A\,|\,\mathscr{G}_Y)$ 则简记为 $\mathbf{E}(X\,|\,Y)$ 与 $\mathbf{P}(A\,|\,Y)$, 并分别称之为 X 关于 Y 的条件数学期望与事件 A 关于 Y 的条件概率. (参阅第二章第 7 节.)

- 同数学分析中一样, 概率论中考察随机变量的各种不同形式的收敛性 (依概率收敛: $X_n \xrightarrow{\mathbf{P}} X$; 几乎必然收敛或几乎处处收敛: $X_n \to X$ ($\mathbf{P}$-a.s.); 依分布收敛: $X_n \xrightarrow{\mathrm{d}} X$ (也可表示为 $X_n \xrightarrow{\mathrm{law}} X$, $\mathrm{Law}\,(X_n) \to \mathrm{Law}\,(X)$, $\mathrm{Law}\,(X_n) \xrightarrow{\mathrm{w}} \mathrm{Law}\,(X)$); p 阶 ($p > 0$) 平均收敛: $X_n \xrightarrow{L^p} X$; 逐点收敛: $X_n(\omega) \to X(\omega)$, $\omega \in \Omega$, $n \geqslant 1$. 参阅第二章第 10 节.)

- 除了随机变量自身的各种收敛性之外, 概率论中还特别关注概率测度, 概率分布以及它们的特征的收敛性.

  最重要的一种收敛形式是: 在某个可测空间 (值得提及的是空间 $\mathbb{R}^n$, $\mathbb{R}^\infty$, $\mathcal{C}$ 与 $\mathcal{D}$) 上给出的概率测度 P_n, $n \geqslant 1$, 弱收敛到测度 P: $P_n \xrightarrow{\mathrm{w}} P$.

  如下的各种结果, 诸如大数律, 中心极限定理, 泊松定理, 收敛到无穷可分分布等等都是关于概率测度弱收敛的结论中的典型例子. (参阅第三章.)

- 概率论中的许多研究都涉及 "概率为 1 地成立" 或 "几乎必然地成立" 的性质. 例如, "0-1 律", 级数的几乎必然收敛, 强大数律, 重对数律 (参阅第四章). 在对它们的研究中, 博雷尔–坎泰利引理起着重要的作用:

  设 $A_1, A_2, \cdots$ 为一序列事件 (集合), $\{A_n \text{ i.o.}\}$ $(\equiv \varlimsup_n A_n \equiv \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k)$ 是那些在序列 $A_1, A_2, \cdots$ 中无限多次 (infinitely often) 出现的 $\omega \in \Omega$ 的集合. 则

  (a) 若 $\sum_{n=1}^{\infty} \mathbf{P}(A_n) < \infty$, 则概率 $\mathbf{P}\{A_n \text{ i.o.}\} = 0$.

  (b) 若 $\sum_{n=1}^{\infty} \mathbf{P}(A_n) = \infty$, 且事件 $A_1, A_2, \cdots$ 相互独立, 则概率 $\mathbf{P}\{A_n \text{ i.o.}\} = 1$.

§4. (狭义) 平稳随机序列

- 在概率空间 $(\Omega, \mathscr{F}, \mathbf{P})$ 上给出的随机变量序列 $X = (X_1, X_2, \cdots)$ 称为狭义平稳的 (严平稳的), 如果其分布律 $\mathrm{Law}\,(X)$(换言之, 分布 P_X) 对每个 $k \geqslant 1$, 都重合于 "移步" 序列 $\theta_k X = (X_{k+1}, X_{k+2}, \cdots)$ 的分布律 $\mathrm{Law}\,(\theta_k X)$.

借助于动力系统理论中的概念, 思想和方法, 容易进行这类序列的性质研究 (见《概率》第二卷第五章).

- 该理论的核心研究内容是保测的可测映射.

  映射 $T: \ \Omega \to \Omega$ 称为可测的, 如果对一切 $A \in \mathscr{F}$, 都有集合 $T^{-1}A = \{\omega : T(\omega) \in A\} \in \mathscr{F}$. 这样的映射称为保测的, 如果对一切 $A \in \mathscr{F}$, 都有

$$\mathbf{P}(T^{-1}A) = \mathbf{P}(A).$$

"狭义平稳序列" 与 "保测变换" 之间的联系清楚地表现如下:

设 T 为某个保测映射, 而 $X_1 = X_1(\omega)$ 为随机变量. $X_n(\omega) = X_1(T^{n-1}\omega)$, 其中 T^{n-1} 是 T 的 $n-1$ 重幂. 则序列 $X = (X_1, X_2, \cdots)$ 是狭义平稳的 (严平稳的).

逆过来的命题也在一定的意义下成立: 对于每个狭义平稳的序列 $X = (X_1, X_2, \cdots)$, 都可以构造一个新的概率空间 $(\widetilde{\Omega}, \widetilde{\mathscr{F}}, \widetilde{\mathbf{P}})$, 一个保持测度 $\widetilde{\mathbf{P}}$ 不变的保测映射 $\widetilde{T}: \ \widetilde{\Omega} \to \widetilde{\Omega}$, 以及一个随机变量 $\widetilde{X}_1 = \widetilde{X}_1(\omega)$, 使得序列 $\widetilde{X} = (\widetilde{X}_1(\omega), \widetilde{X}_1(\widetilde{T}\omega), \cdots)$ 的分布律 $\mathrm{Law}\,(\widetilde{X})$ 重合于 $\mathrm{Law}\,(X)$.

第五章中关于这类性质研究的基本结果有: 自返性 ("庞加莱自返性定理"), 遍历性, 混合性等. 核心结论是伯克霍夫–辛钦遍历性定理, 其 (既是关于保测映射的, 又是关于狭义平稳序列的) 形式之一可陈述如下:

(a) 若 T 为保测遍历变换, $\xi = \xi(\omega)$ 为随机变量, 有 $\mathbf{E}|\xi| < \infty$, 则

$$\lim_n \frac{1}{n} \sum_{k=0}^{n-1} \xi(T^k\omega) = \mathbf{E}\xi \quad (\mathbf{P}\text{-a.s.});$$

(b) 若 $X = (X_1, X_2, \cdots)$ 是狭义平稳的遍历序列, 有 $\mathbf{E}|X_1| < \infty$, 则

$$\lim_n \frac{1}{n} \sum_{k=0}^{n-1} X_k(\omega) = \mathbf{E}X_1.$$

§5. (广义) 平稳随机序列

无论是从理论的观点, 还是从应用的观点, 在研究这样一些序列 X 时, 作如下的假定都是合理的与方便的: 首先假定它们对一切 $n \in \mathbb{Z} = \{0, \pm 1, \pm 2, \cdots\}$ 都有定义; 其次假定序列中的各项都取复数值. 换言之, 假定 $X = (\cdots, X_{-1}, X_0, X_1, \cdots)$, 其中 X_n 都是复值 ($= a_n + ib_n$), 并且对一切 $n \in \mathbb{Z}$, 都有 $\mathbf{E}|X_n|^2 < \infty$; 见第六章第 1 节.

关于 "广义平稳" 概念的基本假定是: 对一切 $n, m \in \mathbb{Z}$, 都有 $\mathbf{E}X_n = \mathbf{E}X_0$ 与 $\mathbf{cov}\,(X_{n+m}, X_m) = \mathbf{cov}\,(X_n, X_0)$.

不失一般性, 可设 $\mathbf{E}X_0 = 0$, 从而 $\mathbf{cov}\,(X_n, X_0) = \mathbf{E}X_nX_0$. 函数 $R(n) = \mathbf{E}X_nX_0$, $n \in \mathbb{Z}$, 称为协方差函数.

- 如下两个结果 (赫格洛茨 (Herglotz) 定理与关于序列 X 的谱表示定理) 表明, 在一定的意义下, 广义平稳序列可以视为在对具有随机振幅的调和分量求和的过程中所产生的随机对象 (带有相应的极限过程, 充满谱频率的所有可能值的区域).

  第一个重要结果是 (第六章第 1 节): 任何协方差函数 $R(n)$, $n \in \mathbb{Z}$ 都具有谱表示:

$$R(n) = \int_{-\pi}^{\pi} e^{i\lambda n} F(d\lambda),$$

其中 $F = F(B)$, $B \in \mathscr{B}([\pi, \pi))$ 是某个 (实值) 有限测度, 而积分理解为勒贝格 - 斯蒂尔切斯积分. 第二个重要结果 (第六章第 3 节) 给出了序列 $X = (X_n)_{n\in\mathbb{Z}}$ 自身的谱表示: 对任何 $n \in \mathbb{Z}$, 都 ($\mathbf{P}$-a.s.) 有

$$X_n = \int_{-\pi}^{\pi} e^{i\lambda n} Z(d\lambda),$$

其中 $Z = Z(\Delta)$, $\Delta \in \mathscr{B}([\pi, \pi))$ 是正交 (复值) 随机测度, 具有性质: $\mathbf{E}Z(\Delta) = 0$, $\mathbf{E}|Z(\Delta)|^2 = F(\Delta)$. (请记住我们的约定: $\mathbf{E}X_0 = 0$.)

  谱函数 $F = F(d\lambda)$ 与谱密度 $f = f(\lambda)$ ($F(B) = \int_B f(\lambda)d\lambda$, $B \in \mathscr{B}([\pi, \pi))$), 如果它们存在, 则在对所考察的过程 X 的相关谱分析中起着重要的作用, 它们给出了关于频率与强度的完全确定的表示, 即关于协方差函数的 “谱成分” 的表示.

  由性质 $\mathbf{E}|Z(\Delta)|^2 = F(\Delta)$ 可以明白谱函数在 “谱随机分量” 的表述中所起的作用, 整体给出了 $X_n = \int_{-\pi}^{\pi} e^{i\lambda n} Z(d\lambda)$, $n \in \mathbb{Z}$.

- 谱特征值的这些性质及其直观性解释了在关于平稳序列 (以及连续时间场合下的平稳随机过程) 的统计学中对它们所表现出的兴趣, 因为统计学就是根据试验数据构造概率 - 统计模型, 运用这些模型去构造关于协方差函数, 谱密度以及它们的特征值的 “好的” 估计量的 (第六章第 4 节).

  还应指出, 正是在所考察的 “广义平稳” 的框架内, (柯尔莫戈洛夫, 维纳) 获得了关于随机序列与过程的过滤结构, 外推和内插方面的周知的结果 (第六章第 6 节).

§6. 鞅

在鞅论的建立过程中, 对概率空间 $(\Omega, \mathscr{F}, \mathbf{P})$ 富有成效地补充了它的结构, 即在其中引入了 σ-代数 (滤子) 流 $(\mathscr{F}_n)_{n\geqslant 0}$, 其直观意义是, $\mathscr{F}_n$ 是到时刻 n 为止 (包括时刻 n) 所观察到的 “信息”. 这样的结构 $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geqslant 0}, \mathbf{P})$ 被形象地称为*滤子概率空间*. 在这样的概率空间中, 还引入了诸如: (关于流 $(\mathscr{F}_n)_{n\geqslant 0}$ 的) “相容性”, “可预报性”, “随机序列”, “鞅性”, “马尔可夫时”, “停时”, 等等概念.

- 与流 $(\mathscr{F}_n)_{n\geqslant 0}$ 相对应的序列 $X = (X_n)_{n\geqslant 0}$, 亦即对每个 n, X_n 是 $\mathscr{F}_n$ 可测的, 称为*鞅*, 如果 $\mathbf{E}|X_n| < \infty$, $n \geqslant 0$, 且 $\mathbf{E}(X_n \,|\, \mathscr{F}_{n-1}) = X_{n-1}$. 当最后一个等式中的等号换为 $\mathbf{E}(X_n \,|\, \mathscr{F}_{n-1}) \geqslant X_{n-1}$ 或 $\mathbf{E}(X_n \,|\, \mathscr{F}_{n-1}) \leqslant X_{n-1}$ 时, 序列 $X = (X_n)_{n\geqslant 0}$ 相应地称为*半鞅* 或*上鞅*.

- 在涉及到一大批饶有兴趣的随机序列类 (例如, 可参阅第七章第 1 节) 的过程中, 鞅论以其丰富的成果而显得非常富足, 在概率论, 数理统计及其应用的数量众多的各个分支之中找到了广泛的应用.

  在这些成果之中首当其冲应当提及的是: 关于在时间的随机变换之下鞅性保持不变的定理 (第 2 节), 关于依概率 1 收敛的定理 (第 4 节), 最大《概率》与 “矩” 不等式 (第 3 节).

  在关于鞅论的应用方面, 有第 10 ∼ 13 节中的材料为例, 其中涉及了诸如: 保险业破产概率问题、“随机金融数学” 、以及随机序列的 “最佳停时” 问题等等.

§7. 马尔可夫链

除了一系列已经在第八章正文中出现过的关于马尔可夫链的重要概念和符号之外, 本节还将引入一些新的概念与符号, 因为它们出现在了一些新增加的题目中.

- 如同鞅论一样, 我们所考察的取值于某个相空间 $(E, \mathscr{E})$ 的 (广义) 马尔可夫链 $X = (X_n)_{n\geqslant 0}$ 是定义在某个滤子概率空间 $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geqslant 0}, \mathbf{P})$ 上的. 此处, 对每个 $n \geqslant 0$, 都假定 X_n 为 $\mathscr{F}_n/\mathscr{E}$ 可测的. 刻画我们所考察的序列的最基本的假定是: 我们的序列具有*广义马尔可夫性*, 即对一切 $n \geqslant 0$ 与 $B \in \mathscr{E}$, 都 ($\mathbf{P}$-a.s.) 有

$$\mathbf{P}(X_{n+1} \in B \,|\, \mathscr{F}_n)(\omega) = \mathbf{P}(X_{n+1} \in B \,|\, X_n)(\omega)$$

(往往将 $\mathbf{P}(X_{n+1} \in B \,|\, X_n)(\omega)$ 写为 $\mathbf{P}(X_{n+1} \in B \,|\, X_n(\omega))$).

  若 σ-代数 $\mathscr{F}_n = \mathscr{F}_n^X \equiv \sigma(X_0, X_1, \cdots, X_n)$, 则所考察的性质为*狭义马尔可夫性*, 并说 $X = (X_n)_{n\geqslant 0}$ 是*马尔可夫链*.

  若相空间 $(E, \mathscr{E})$ 是博雷尔的, 则根据第二章第 7 节定理 5, 存在*正则的条件分布* $P_n(x; B)$, 使得 ($\mathbf{P}$-a.s.地) 有

$$\mathbf{P}(X_n \in B | X_{n-1}(\omega)) = P_n(X_{n-1}(\omega); B), \quad B \in \mathscr{E}, \quad n \geqslant 1.$$

  函数 $P_n(x; B)$ 在马尔可夫理论中通常称为 (由 E 到 $\mathscr{E}$ 的) *转移函数*, 或*马尔可夫核*. 若函数 $P_n(x; B)$, $n \geqslant 1$, 与 n 无关 ($= P(x; B)$), 则相应的 (*广义的或狭义的*) 马尔可夫链称为*齐次的*.

除了转移函数之外, 刻画马尔可夫链的另一个重要特征是初始分布 $\pi = \pi(B)$, $B \in \mathscr{E}$:

$$\pi(B) = \mathbf{P}\{X_0 \in B\}.$$

$(\pi, P_1, P_2, \cdots)$(在齐次场合, (π, P)) 完全确定了马尔可夫序列 $X = (X_0, X_1, \cdots)$ 的概率分布.

- 在第八章关于马尔可夫链理论的论述中, 我们根据当代的习惯, 数次改变观点, 认为起始的对象不是 $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geqslant 0}, \mathbf{P})$, 以及取值于 E 的 $\mathscr{F}/\mathscr{E}$ 可测的随机变量 X_n 的序列 $X = (X_0, X_1, \cdots)$, 其中 $(E, \mathscr{E})$ 是相空间, 而是由 E 向 $\mathscr{E}$ 的 "转移函数" 组 $(P_1, P_2, \cdots)$, 其中 $P_n = P_n(x; B)$, $x \in E$, $B \in \mathscr{E}$, 而 $(E, \mathscr{E})$ 是某个给定的相空间. (在齐次场合只有一个转移函数 $P = P(x; B)$.) 因此, 成组地构造出 (例如, 根据依奥涅斯库 - 图尔其 (Ионеску Тулчи) 定理, 其中取坐标空间 $(E^\infty, \mathscr{E}^\infty)$ 作为 $(\Omega, \mathscr{F})$) 概率测度族 $\{P_x, x \in E\}$, 标准地形成给定序列 $X = (X_0, X_1, \cdots)$(若 $\omega = (x_0, x_1, \cdots)$, 则 $X_n(\omega) = x_n$), 它对于每个 $x \in E$, 关于 P_x 为马尔可夫的, 此时 $P_x\{X_0 = x\} = 1$. 如果 $\pi = \pi(B)$, $B \in \mathscr{E}$, 是某个 "初始" 分布, 则以 P_π 表示 $(E^\infty, \mathscr{E}^\infty)$ 上的按如下方式构造的分布: $P_\pi(A) = \int_E P_x(A)\pi(dx)$, $A \in \mathscr{E}^\infty$.

  关于 P_x, 序列 X 自然称为由点 x 出发的马尔可夫链 (即有 $X_0(\omega) = x$, (P_x-a.s.)). 关于测度 P_π, 序列 X 则称为依 π 为初始分布的马尔可夫链 (亦即 $P_\pi\{X_0 \in B\} = \pi(B)$, $B \in \mathscr{E}$).

- 为了陈述两个 (新的) 马尔可夫性质: 广义马尔可夫性与严马尔可夫性, 有益的做法是引入移步算子 θ 和它的 "幂" θ_n 与 θ_τ (其中, τ 为马尔可夫时).

  算子 $\theta: \Omega \to \Omega$ 称为移步算子, 如果对每个 $\omega = (x_0, x_1, \cdots)$, 都有

$$\theta(\omega) = (x_1, x_2, \cdots).$$

换言之, 轨道 $(x_0, x_1, \cdots)$ 在算子 θ 的作用下, "移步" 为轨道 $(x_1, x_2, \cdots)$. (在第五章中, 在讨论狭义平稳序列及其相应的动力系统时, 也曾引入过 Ω 到自身的映射, 当时是用 T 表示的.)

以 $\theta_0 = I$ 表示单位的恒等的变换 $(\theta_0(\omega) = \omega)$, 对 $n \geqslant 1$, 可以将算子 θ 的 n 次幂 θ_n 定义为: $\theta_n = \theta_{n-1} \circ \theta (= \theta \circ \theta_{n-1})$, 亦即 $\theta_n(\omega) = \theta_{n-1}(\theta(\omega))$.

若 τ 为马尔可夫时 $(\tau \leqslant \infty)$, 则以 θ_τ 表示这样的算子: 它仅作用在集合 $\Omega_\tau = \{\omega: \tau(\omega) < \infty\}$ 上, 且当 $\tau = n$ 时, 有 $\theta_\tau = \theta_n$, 亦即对一切使得 $\tau(\omega) = n$ 的 ω, 都有

$$\theta_\tau(\omega) = \theta_n(\omega).$$

若 $H = H(\omega)$ 为 $\mathscr{F}$ 可测的函数 (例如 $\tau = \tau(\omega)$, $X_m = X_m(\omega)$), 则用 $H \circ \theta_n$ 表示函数 $(H \circ \theta_n)(\omega) = H(\theta_n(\omega))$.

若 σ 为马尔可夫时, 则 $H \circ \theta_\sigma$ 仅定义在集合 $\Omega_\sigma = \{\omega: \sigma(\omega) < \infty\}$ 上, 且当 $\sigma(\omega) = n$ 时, 有 $H \circ \theta_\sigma = H \circ \theta_n$, 亦即若 $\omega \in \{\sigma(\omega) = n\}$, 则 $(H \circ \theta_\sigma)(\omega) = (H \circ \theta_n)(\omega) = H(\theta_n(\omega))$.

由这些定义, 可以特别地推出

$$X_m \circ \theta_n = X_{m+n},$$
$$X_m \circ \theta_\sigma = X_{m+\sigma} \quad \text{在集合 } \Omega_\sigma \text{ 上},$$

而若 τ 与 σ 都是有限马尔可夫时, 则

$$X_\tau \circ \theta_\sigma = X_{\tau \circ \theta_\sigma + \sigma}.$$

与算子 $\theta_n: \Omega \to \Omega$ 相对应, 可以定义逆算子 $\theta_n^{-1}: \mathscr{F} \to \mathscr{F}$ 为: 若 $A \in \mathscr{F}$, 则

$$\theta_n^{-1}(A) = \{\omega: \theta_n(\omega) \in A\}.$$

若将 A 取为集合 $\{\omega: X_m(\omega) \in B\}$, $B \in \mathscr{E}$, 则有

$$\theta_n^{-1}(A) = \{\omega: X_{m+n}(\omega) \in B\},$$

亦即

$$\theta_n^{-1}(X_m^{-1}(B)) = X_{m+n}^{-1}(B).$$

(在第八章第 2 节中, 有许多关于算子 $\theta_n, \theta_\sigma, \theta_n^{-1}$ 等等的性质的新的习题.)

- 上面所引入的算子 θ_n 可以用来证明 (第 2 节定理 1) 所谓的*推广的马尔可夫性*: 若 $H = H(\omega)$ 为有界的 (或非负的)$\mathscr{F}$ 可测函数, 则对任何初始分布 π 与任何 $n \geqslant 0$, 都有

$$E_\pi(H \circ \theta_n \,|\, \mathscr{F}_n^X)(\omega) = E_{X_n(\omega)} H \quad (P_\pi\text{-a.s.}),$$

其中, E_π 表示对测度 π 求均值, 而 $E_{X_n(\omega)}H$ 则应理解为先形成函数 $\psi(x) = E_x H$, 再根据定义, 将 $E_{X_n(\omega)}H$ 理解为 $\psi(X_n(\omega))$.

这一推广了的马尔可夫性还可以进一步推广, 包括可将其中确定的时刻 n 换为有限的马尔可夫时 τ. 确切地说, 我们有: 若 $(H_n)_{n\geqslant 0}$ 为一序列有界的 (或非负的)$\mathscr{F}$ 可测函数, τ 为有限的马尔可夫时, 则由马尔可夫性可推出所谓的*严马尔可夫性*, 其中包括: 对每个初始分布 π, 都有

$$E_\pi(H_\tau \circ \theta_\tau \,|\, \mathscr{F}_\tau^X)(\omega) = \psi(\tau(\omega), X_{\tau(\omega)}(\omega)) \quad (P_\pi\text{-a.s.}),$$

其中, $\psi(n, x) = E_x H_n$.

请注意, 应当对 $H_\tau \circ \theta_\tau = (H_\tau \circ \theta_\tau)(\omega)$ 作这样的理解: 当 $\tau(\omega) = n$ 时, 有 $(H_\tau \circ \theta_\tau)(\omega) = (H_n \circ \theta_n)(\omega)$.

- 正如上面所已经指出的, 在现代理论中, 在某个相空间 $(E, \mathscr{E})$ 中取值的齐次马尔可夫链 $X=(X_n)_{n\geqslant 0}$ 被其初始分布 $\pi=\pi(dx)$ 与转移函数 $P=P(x;B)$, $x\in E$, $B\in\mathscr{E}$, 所完全决定. 在此, 空间 $(E^\infty, \mathscr{E}^\infty)$ 中的概率分布 P_x 仅由转移函数 $P=P(x;B)$ 决定.

值得注意的是, 转移函数 (或称为马尔可夫核) 还是数学分析中的另一个分支 —— 所谓位势理论的基础. 因此, 毫不奇怪, 该理论与齐次马尔可夫链有着非比寻常的密切关系, 而且这种联系形成了相互促进的局面.

现在来谈几个无论对于位势理论, 还是对于马尔可夫理论, 都是重要的与必须的问题.

与转移函数 $P=P(x;B)$, $x\in E$, $B\in\mathscr{E}$ 有关, 作用在函数 $g=g(x)$ 上的 (一步) 转移线性算子为

$$\mathbb{P}g(x)=\int_E g(y)P(x;dy).$$

(常常写为 $(\mathbb{P}g)(x)$.) 作为算子 $\mathbb{P}$ 的定义域, 人们考察了那些对一切 $x\in E$, 使得积分 $\int_E g(y)P(x;dy)$ 确定的 $\mathscr{E}$ 可测的函数 $g=g(x)$. 该积分对于非负函数类 (记为 $\mathscr{E}_+$) 或有界函数类 (记为 $b\mathscr{E}$) 是确定的.

以 $\mathbb{I}$ 表示单位 (恒等) 算子 ($\mathbb{I}g(x)=g(x)$), 我们来引入算子 $\mathbb{P}_n$ (n 步转移): 对 $n\geqslant 1$, 令 $\mathbb{P}_n=\mathbb{P}(\mathbb{P}_{n-1})$, 或等价地, $\mathbb{P}_n=\mathbb{P}_{n-1}(\mathbb{P})$, 其中 $\mathbb{P}_0=\mathbb{I}$.

易知, 对于一切使得积分 $\int_E g(y)P^n(x;dy)$ 确定的 $\mathscr{E}$ 可测的函数 $g=g(x)$, 都有

$$\mathbb{P}_n g(x)=\mathbf{E}_x g(X_n),$$

其中 $P^n=P^n(x;dy)$ 为 n 步转移概率 (参阅第八章第 1 节).

若 τ 为 (关于流 $(\mathscr{F}_n^X)_{n\geqslant 0}$, 其中 $\mathscr{F}_n^X=\sigma(X_0,X_1,\cdots,X_n)$ 的) 马尔可夫时, 则用 $\mathbb{P}_\tau$ 表示按如下方式作用在函数 $g=g(x)$ 上的算子:

$$\mathbb{P}_\tau g(x)=\mathbf{E}_x[I(\tau<\infty)g(X_\tau)].$$

我们注意到, 若 $g(x)\equiv 1$, 则

$$\mathbb{P}_\tau 1(x)=P_x\{\tau<\infty\}.$$

用算子 $\mathbb{P}_n$, $n\geqslant 0$, 可以构造 (一般说来, 是非有界的) 算子

$$\mathbb{U}=\sum_{n\geqslant 0}\mathbb{P}_n,$$

称为算子 $\mathbb{P}_n$ (或相应的马尔可夫链) 的位势.

若 $g\in\mathscr{E}_+$, 则

$$\mathbb{U}g=\sum_{n\geqslant 0}\mathbb{P}_n g=(\mathbb{I}+\mathbb{P}\mathbb{U})g,$$

或简单地写为

$$\mathbb{U}=\mathbb{I}+\mathbb{P}\mathbb{U}.$$

函数 $\mathbb{U}g$ 称为函数 g 的位势.

若 $g(x)=I_B(x)$ 是集合 $B\in\mathscr{E}$ 的示性函数, 则

$$\mathbb{U}I_B(x)=\sum_{n\geqslant 0}E_xI_B(X_n)=E_xN_B,$$

其中 N_B 是访问集合 B 的次数. 当 $x\in E$ 固定时, 函数 $U(x,B)=\mathbb{U}I_B(x)$ 是在集合 $B\in\mathscr{E}$ 上的测度. 有时也将这一测度称为位势 - 测度. 若 $B=\{y\}$ 为单点集, $y\in E$, 则将函数 $U(x,\{y\})$ 记为 $G(x,y)$, 称为 (算子 $\mathbb{P}$ 或相应的马尔可夫链的) 格林函数. 格林函数的直观意义是清楚的: $G(x,y)=E_xN_{\{y\}}$ 就是在假设 $X_0=x$ 之下访问状态 y 的平均访问次数.

与由算子 $\mathbb{P}$ 得到位势 $\mathbb{U}$ 的做法类似, 也可以按如下公式由转移函数 (马尔可夫核)$P=P(x;B)$ 得到核 $Q=Q(x;B)$:

$$Q(x;B)=\sum_{n\geqslant 0}P^n(x;B)\quad (=(I_B(x)+PQ(x;B)).$$

由于 $\mathbb{P}_nI_B(x)=P^n(x;B)$, 故显然有 $U(x;B)=Q(x;B)$.

- 与算子 $\mathbb{P}$ 有关的还有一个重要算子

$$\mathbb{L}=\mathbb{P}-\mathbb{I},$$

其中, $\mathbb{I}$ 是单位 (恒等) 算子. 在马尔可夫理论中, 算子 $\mathbb{L}$ 称为 (具有转移函数 $P=P(x;B)$ 的齐次马尔可夫链的)生成算子. 算子 $\mathbb{L}$ 的定义域 $\mathscr{D}_{\mathbb{L}}$ 是使得 $\mathbb{P}g-g$ 确定的所有 $\mathscr{E}$ 可测的函数 $g=g(x)$ 构成的集合.

若 $h\in\overline{\mathscr{E}}_+$(即 h 是 $\mathscr{E}$ 可测的非负函数, 取值于 $\overline{\mathbb{R}}_+$), 则它的位势 $H=\mathbb{U}h$ 满足关系式:

$$H=h+\mathbb{P}H,$$

(此因 $\mathbb{U}=\mathbb{I}+\mathbb{P}\mathbb{U}$). 事实上, 若 $h\in\mathscr{D}_{\mathbb{L}}$, 则 H 是泊松方程

$$\mathbb{L}V=-h$$

(在函数类 $V\in\mathscr{D}_{\mathbb{L}}$ 中的) 解.

若方程 $V=h+\mathbb{P}V$ 还有解 $W\in\overline{\mathscr{E}}_+$(或方程 $\mathbb{L}V=-h$ 还有解, 如果 $W\in\mathscr{D}_{\mathbb{L}}$), 则由于 $W=h+\mathbb{P}W\geqslant h$, 故由归纳法可知, 对任何 $n\geqslant 1$, 都有 $W\geqslant\sum_{k=1}^{n}\mathbb{P}_kh$, 从而就有 $W\geqslant H$. 如此一来, 在函数类 $\overline{\mathscr{E}}_+$ 中, 位势 $H=\mathbb{U}h$ 就是方程组 $V=h+\mathbb{P}V$ 的最小解. 请记住 $\mathbb{U}h(x)=\mathbf{E}_x\sum_{k=0}^{\infty}h(X_k)$.

- 设函数 $f = f(x)$, $x \in E$ 属于类 $\overline{\mathscr{E}}_+$, 称 f 对于算子 $\mathbb{P}$ 是峰态的(或对于具有转移函数 $P = P(x;\, B)$ 的马尔可夫链是峰态的, 或 $\mathbb{P}$-峰态的), 如果

$$\mathbb{P}f \leqslant f$$

(或写为 $\mathbf{E}_x f(X_1) \leqslant f(x),\ x \in E$).

根据这一定义, 函数 $h \in \overline{\mathscr{E}}_+$ 的位势 $H = \mathbb{U}h$ 是峰态的函数.

类 $\overline{\mathscr{E}}_+$ 中的函数 $f = f(x)$ 称为调和的 (或不变的), 如果

$$\mathbb{P}f = f$$

(亦即 $\mathbf{E}_x f(X_1) = f(x),\ x \in E$).

在位势理论 (其中的概念之一是峰态性) 与概率论 (确切地说, 鞅论) 之间的联系之一可以直观地表述为如下命题: 如果 $X = (X_n)_{n\geqslant 0}$ 是马尔可夫链, 具有初始分布 π 以及 $(E^\infty, \mathscr{E}^\infty)$ 中由分布 P_π 生成的转移函数 $P = P(x;\, B)$ 的马尔可夫链, 而 $f = f(x)$ 是 $\mathbb{P}$-峰态的函数, 则序列 $Y_n = f(X_n)$, $Y = (Y_n, \mathscr{F}_n^X, P_\pi)_{n\geqslant 0}$ 是 "非负定上鞅" 序列:

$$Y_n \text{ 为 } \mathscr{F}_n^X\text{-可测}, \quad n \geqslant 0,$$

$$E_\pi(Y_{n+1} | \mathscr{F}_n^X) \leqslant Y_n \quad P_\pi\text{-a.s.} \quad n \geqslant 0.$$

如果还有 $\mathbf{E}\, Y_n < \infty$, $n \geqslant 0$, 那么该序列就是 (通常意义下的) 上鞅.

有趣的是, 这样的 "非负定上鞅" 序列 Y 保持了非负上鞅的性质 (参阅第七章第 4 节定理 1): 概率 P_π 为 1 地存在极限 $\lim\limits_n Y_n$ $(= Y_\infty)$; 且若 $P_\pi\{Y_0 < \infty\} = 1$, 则 $P_\pi\{Y_\infty < \infty\} = 1$. (该结论的证明见第七章第 4 节第 24 题.)

- (属于类 $\mathscr{E}_+$ 或 $\overline{\mathscr{E}}_+$ 的) 非负函数 $h = h(x)$ 的位势 $H(x) = \mathbb{U}h(x)$ 满足关系式 $H(x) = h(x) + \mathbb{P}H(x)$, 由此可以推出

$$H(x) \geqslant \max\,(h(x),\ \mathbb{P}H(x)), \quad x \in E.$$

这表明, 函数 $h(x)$ 的位势 $H(x)$, 一方面控制着该函数 (意即 $H(x) \geqslant h(x)$, $x \in E$), 另一方面, 还是峰态的. 换句话说, 函数 $h(x)$ 的位势 $H(x)$ 是其峰态控制函数的例子. 在许多场合下 (例如在考察最优停时的问题中, 参阅第八章第 9 节), 存在着寻找关于 $\mathscr{E}$-可测的非负函数 $g = g(x)$ 的最小峰态控制函数的问题.

位势理论以下述方式回答了这个问题.

引入算子 $\mathbb{Q}$ 以下述方式作用在函数 $g = g(x)$ 上:

$$\mathbb{Q}g(x) = \max\,(g(x),\, \mathbb{P}g(x)).$$

函数 $g(x)$ 的最小峰态控制函数 $s(x)$ 即可按照下述公式求出:

$$s(x) = \lim_n \mathbb{Q}^n g(x),$$

在此有

$$s(x) = \max\,(g(x),\, \mathbb{P}s(x)), \quad x \in E.$$

由这一等式, 特别地, 可知: 若 $s \in \mathscr{D}_{\mathbb{L}}$, 则

$$\mathbb{L}s(x) = 0, \quad x \in C_g,$$

$$s(x) = g(x), \quad x \in D_g,$$

其中, $C_g = \{x:\ s(x) > g(x)\}$, $D_g = E\backslash C_g$. (该证明可见第八章第 9 节, 在那里, $\mathbb{P}$ 用 T 表示, 而 $\mathbb{Q}$ 则写为 Q.)

- 位势理论对关于算子 $\mathbb{P}$ 的狄利克雷问题的解的刻画问题给予了极大的关注: (从某个函数类, 例如, $\overline{\mathscr{E}}_+$, $\mathscr{E}_+$, $b\mathscr{E}$, 等等之中) 寻找非负函数 $V = V(x)$, $x \in E$, 使得

$$V(x) = \begin{cases} \mathbb{P}V(x) + h(x), & x \in C, \\ g(x), & x \in D. \end{cases}$$

其中 C 是 E 中的某个给定的子集 (通常称为 "区域"), $D = E\backslash C$, 而 h 与 g 亦为给定的 $\mathscr{E}$ 可测的非负函数.

若仅考察属于 $\mathscr{D}_{\mathbb{L}}$ 的解 V, 则所列的方程组等价于

$$\mathbb{L}V(x) = -h(x), \qquad x \in C,$$

$$V(x) = g(x), \qquad x \in D.$$

方程 $\mathbb{L}V = -h$ 称为区域 C 中的泊松方程, 并且通常说到狄利克雷问题, 总是意味着, 在区域 C 中求这个方程解, 而在区域 D 中则等于预先给定的函数 g.

令人印象深刻的是, 所陈述的 (非概率的) 狄利克雷问题, 在考虑了具有转移函数 $P = P(x;\, B)$ 的马尔可夫链, 并根据其构造出算子 $\mathbb{P}$ 之后, 得以顺利解决.

设 $X = (X_n)_{n\geqslant 0}$ 为这样的马尔可夫链, $\tau(D) = \inf\{n \geqslant 0:\ X_n \in D\}$. (如同通常, 若 $\{\cdot\} = \varnothing$, 则 $\tau(D) = \infty$.)

可以断言, 若 h 与 g 属于类 $\overline{\mathscr{E}}_+$, 则狄利克雷问题的解存在, 并且最小非负解 $V_D(x)$ 由下式给出:

$$V_D(x) = \mathbf{E}_x[I(\tau(D) < \infty) g(X_{\tau(D)})] + I_C(x)\mathbf{E}_x\left[\sum_{k=0}^{\tau(D)-1} h(X_k)\right].$$

(关于这一结果的证明的提示参阅第八章第 8 节第 11 题.)

我们指出若干特殊情形:

(a) 若 $h=0$, 则所找的函数 $V=V(x)$ 在区域 C 中是调和的, 在区域 D 中等于函数 g, 故最小非负解 $V_D(x)$ 由下式给出:

$$V_D(x)=\mathbf{E}_x[I(\tau(D)<\infty)g(X_{\tau(D)})].$$

特别地, 若 $g(x)\equiv 1,\ x\in D$, 则

$$V_D(x)=\mathbf{P}_x\{\tau(D)<\infty\}.$$

事实上, 在由方程状态 $X_0=x$ 出发的假设下, 迟早到达集合 D 的概率 $\mathbf{P}_x\{\tau(D)<\infty\}$ (在区域 C 内) 是调和函数. 易知, 若 $x\in D$, 则 $\mathbf{P}_x\{\tau(D)<\infty\}=1$, 因为此时 $\tau(D)=0$.

(b) 若 $g(x)=0,\ x\in D$, 且 $h(x)=1,\ x\in C$, 则所考察的方程组为

$$V(x)=\begin{cases}\mathbb{P}V(x)+1, & x\in C,\\ 0, & x\in D,\end{cases}$$

此时可得其最小非负解为

$$V_D(x)=I_C(x)\mathbf{E}_x\left[\sum_{k=0}^{\tau(D)-1}1\right]=\begin{cases}\mathbf{E}_x\tau(D), & x\in C,\\ 0, & x\in D.\end{cases}$$

如此一来, 首次到达集合 D 的时刻的数学期望 $\mathbf{E}_x\tau(D)$ 就是所列方程组的最小非负解.

- 在描述相空间 $(E,\mathscr{E})$ 中的随机游动的马尔可夫类序列中, 在

$$E=\mathbb{Z}^d=\{0,\pm1,\pm2,\cdots\}^d$$

中的简单对称随机游动, 亦即沿着整数格子点 $\mathbb{Z}^d$ 的随机游动, 起着特别的作用, 其中 $d=1,2,\cdots$(参阅第八章第 8 节). 这种 (在 “整个” 空间 $E=\mathbb{Z}^d$ 中的) 随机游动 $X=(X_n)_{n\geqslant 0}$ 可以构造性地给出, 令

$$X_n=x+\xi_1+\cdots+\xi_n,$$

其中, 在某个概率空间 $(\Omega,\mathscr{F},\mathbf{P})$ 上给出的 d 维随机向量 $\xi_1,\xi_2,\cdots$ 是独立同分布的, 有

$$\mathbf{P}\{\xi_1=e\}=(2d)^{-1}$$

(其中 $e=(e_1,\cdots,e_d)$ 是空间 $\mathbb{R}^d$ 中的标准基单位向量, 亦即 $e_i=0,-1$ 或 1, 且 $\|e\|=|e_1|+\cdots+|e_d|=1$). 这种始于点 $x\in\mathbb{Z}^d$ 的随机游动 $X=(X_n)_{n\geqslant 0}$ 的特征是, 游动的 “质点” 自每个状态均等概率地转移到 $2d$ 个最近邻的整点之上.

此时相应的算子 $\mathbb{P}$ 的构造非常简单:

$$\mathbb{P}f(x)=\mathbf{E}_xf(x+\xi_1)=\frac{1}{2d}\sum_{\|e\|=1}f(x+e).$$

此时将相应的算子 $\mathbb{L}=\mathbb{P}-\mathbb{I}$ 称为 (离散的)拉普拉斯算子, 并记为 Δ, 易见有

$$\Delta f(x)=\frac{1}{2d}\sum_{\|e\|=1}(f(x+e)-f(x)).$$

上面所说到的狄利克雷问题对于所考察的简单随机游动自然会有某些变化, 值得注意的是, 离开区域 $C\subseteq\mathbb{Z}^d$ 时只能落在边界集合

$$\partial C=\{x:\ x\in\mathbb{Z}^d,\ x\notin C \text{ 且对某个 } y\in C,\ \text{有 } \|x-y\|=1\}.$$

这种情况导致 (关于泊松方程) 的狄利克雷问题此时的标准形式为: 对于给定的区域 $C\subseteq\mathbb{Z}^d$ 和给定的函数 $h=h(x),\ x\in C$ 与 $g=g(x),\ x\in\partial C$, 寻找函数 $V=V(x)$, 使得

$$\begin{aligned}\Delta V(x)&=-h(x), & x&\in C,\\ V(x)&=g(x), & x&\in\partial C.\end{aligned}$$

若区域 C 为有限个点构成的, 则对一切 $x\in C$, 都有 $\mathbf{P}_x\{\tau(\partial C)<\infty\}=1$, 其中 $\tau(\partial C)=\inf\{n\geqslant 0:\ X_n\in\partial C\}$(参阅第八章第 8 节第 12 题). 这使得可用上面所说到的方法证明, 所考察的问题的唯一解, 由下式给出: 对 $x\in C\cup\partial C$, 有

$$V_{\partial C}(x)=\mathbf{E}_xg(X_{\tau(\partial C)})+I_C(x)\mathbf{E}_x\left[\sum_{k=0}^{\tau(\partial C)-1}h(X_k)\right].$$

(由于区域 C 为有限的, 故无需假定 $h(x)$ 与 $g(x)$ 为非负函数.)

特别地, 若 $h=0$, 则在区域 C 中调和, 且当 $x\in\partial C$ 时等于 $g(x)$ 的唯一函数是:

$$V_{\partial C}(x)=\mathbf{E}_xg(X_{\tau(\partial C)}).$$

下面来引述关于如下的齐次狄利克雷问题的一些结果: 在非有界区域 C 中,

$$\begin{aligned}\Delta V(x)&=0, & x&\in C,\\ V(x)&=g(x), & x&\in\partial C.\end{aligned}$$

若 $d\leqslant 2$, 则根据波利亚定理 (第八章第 8 节), 有 $\mathbf{P}_x\{\tau(\partial C)<\infty\}=1$, 并且 C 为有限区域时的方法仍适用于此, 故可得到如下结果: 当 $g=g(x)$ 为有界函数时, 则在 $C\cup\partial C$ 中为有界类的解集中, 存在着唯一的解, 其形式与上面的相同, 即:

$$V_{\partial C}(x)=\mathbf{E}_xg(X_{\tau(\partial C)}).$$

必须指出, 即使在 $g = g(x)$ 为有界函数时, 我们的问题的解也可能为 (不是一个) 非有界解. 典型的例子如下.

设 $d = 1$, $C = \mathbb{Z}\backslash\{0\}$, 意即 $\partial C = \{0\}$. 令 $g(0) = 0$, 则易见, 任何无界函数 $V(x) = \alpha x$, $\alpha \in \mathbb{R}$, 都是狄利克雷问题 ($\Delta V(x) = 0$, $x \in \mathbb{Z}\backslash\{0\}$, 且 $V(0) = g(0)$) 的解.

若 $d \geqslant 3$, 则狄利克雷问题 ($\Delta V(x) = 0$, $x \in C$, 且 $V(x) = g(x)$, $x \in \partial C$) 的有界函数解的存在性与唯一性问题, 本质上取决于是否对一切 $x \in C$, 都有 $\mathbf{P}_x\{\tau(\partial C) < \infty\} = 1$. 若该条件满足, 则在有界函数类中, 解存在且唯一, 具有形式: 对一切 $x \in C \cup \partial C$, 有 $V_{\partial C}(x) = \mathbf{E}_x g(X_{\tau(\partial C)})$.

而若条件 $\mathbf{P}_x\{\tau(\partial C) < \infty\} = 1$, $x \in C$, 不满足, 则 (在 $g = g(x)$, $x \in \partial C$, 为有界函数的场合下) 狄利克雷问题 ($\Delta V(x) = 0$, $x \in C$, 且 $V(x) = g(x)$, $x \in \partial C$) 的所有有界解均具有如下形式:

$$V_{\partial C}^{(\alpha)}(x) = \mathbf{E}_x[I(\tau(\partial C) < \infty)g(X_{\tau(\partial C)})] + \alpha \mathbf{P}_x\{\tau(\partial C) = \infty\},$$

其中, $\alpha \in \mathcal{R}$. (参阅, 例如, [75] 定理 1.4.9.)

- 在 (第八章) 阐述具有可数状态集合的马尔可夫链的分类问题时, 我们摆脱了 30 年代 (柯尔莫戈洛夫, 弗雷歇, 道博林等) 的那种最为复杂的局面, 一种自然的分类方式, 一方面应当由其转移概率矩阵的代数性质所决定, 另一方面, 也取决于其转移概率随时间增长时的渐近性状.

从那时起, 由其转移概率的代数性质所决定的如下的一些概念:

本质与非本质状态,
可达状态与连通状态,
不可约类与循环类,

与由其转移概率的极限性状所决定的如下的一些概念:

自返性与非自返性,
正定状态与零状态,
不变 (平稳) 分布,
遍历分布与遍历定理,

成为了马尔可夫链理论中所集中研究的标准概念.

现在已经越来越清楚, 关于渐近性质的研究在与位势理论的概念相结合后更为方便, 对于后者的基本概念 (位势, 调和函数, 峰态函数, $\cdots$) 前面已做了介绍.

从第八章的阐述中可见, 在对于马尔可夫链极限性状的研究中, 所谓 "再生循环" 法已经逐渐成为最基本的工具, 作为该方法基础的是如下的观察.

设 x 是 E 中的某个状态. 定义再生马尔可夫时序列 $(\sigma_x^k)_{k\geqslant 0}$ 如下: 令 $\sigma_x^0 = 0$, $\sigma_x^1 = \sigma_x$, 其中

$$\sigma_x = \inf\{n > 0: X_n = x\},$$

再归纳式地, 对 $k \geqslant 2$, 在 $\sigma_x^{k-1} < \infty$ 的集合上, 令

$$\sigma_x^k = \inf\{n > \sigma_x^{k-1}: X_n = x\}.$$

换言之, 即有

$$\sigma_x^k = \begin{cases} \sigma_x^{k-1} + \sigma_x \circ \theta_{\sigma_x^{k-1}}, & \text{如果 } \sigma_x^{k-1} < \infty, \\ \infty, & \text{如果 } \sigma_x^{k-1} = \infty. \end{cases}$$

如下的性质说明了名称 "再生时刻" 与术语 "再生循环" 的含义:

1) 在集合 $\{\sigma_x^k < \infty\}$ 上, 有 $X_{\sigma_x^k} = x$;

2) 在集合 $\{\sigma_x^k < \infty\}$ 上, 序列 $(X_{\sigma_x^k+n})_{n\geqslant 0}$(关于测度 $\mathbf{P}_x$) 与 $(X_0, X_1, \cdots, X_{\sigma_x^k-1})$ 独立;

3) 若对一切 $\omega \in E^\infty$ 都有 $\sigma_x^k(\omega) < \infty$, 则序列 $(X_{\sigma_x^k+n})_{n\geqslant 0}$ 的 $\mathbf{P}_x$-分布与序列 $(X_n)_{n\geqslant 0}$ 的分布相同;

4) 若对一切 $\omega \in E^\infty$ 都有 $\sigma_x^k(\omega) < \infty$, 则 "再生循环"

$$(X_0, X_1, \cdots, X_{\sigma_x^1-1}), \cdots, (X_{\sigma_x^{k-1}}, X_{\sigma_x^{k-1}+1}, \cdots, X_{\sigma_x^k-1})$$

是 $\mathbf{P}_x$-独立的;

5) $\mathbf{P}_x\{\sigma_x^k < \infty\} = \mathbf{P}_x\{\sigma_x^{k-1} < \infty\}\mathbf{P}_x\{\sigma_x < \infty\}$, 因而 $\mathbf{P}_x\{\sigma_x^n < \infty\} = [\mathbf{P}_x\{\sigma_x < \infty\}]^n$;

6) 若 $N_x = \sum_{n\geqslant 0} I_{\{x\}}(X_n)$ ($= N_{\{x\}}$ 即处于状态 x 的时刻数目), 则平均数目 $\mathbf{E}_x N_x$ ($= \mathbf{E}_x N_{\{x\}} = G(x, x)$) 满足下式:

$$\mathbf{E}_x N_x = 1 + \sum_{n\geqslant 1} \mathbf{P}_x\{\sigma_x^n < \infty\} = 1 + \sum_{n\geqslant 1} [\mathbf{P}_x\{\sigma_x < \infty\}]^n;$$

7) 由上式推出

$$\mathbf{P}_x\{\sigma_x < \infty\} = 1 \quad \Leftrightarrow \quad \mathbf{E}_x N_x = \infty \quad \Leftrightarrow \quad \mathbf{P}_x\{N_x = \infty\} = 1,$$
$$\mathbf{P}_x\{\sigma_x < \infty\} < 1 \quad \Leftrightarrow \quad \mathbf{E}_x N_x < \infty \quad \Leftrightarrow \quad \mathbf{P}_x\{N_x < \infty\} = 1;$$

8) 对于 $y \neq x$, 有

$$G(x, y) = \mathbf{P}_x\{\sigma_y < \infty\}G(y, y),$$

亦即有

$$\mathbf{E}_x N_y = \mathbf{P}_x\{\sigma_y < \infty\}\mathbf{E}_y N_y;$$

9) 若对一切 $k \geqslant 1$, 都有 $\mathbf{P}_x\{\sigma_x < \infty\} = 1$, 则由再生区间长度构成的序列 $(\sigma_x^k - \sigma_x^{k-1})_{k\geqslant 0}$ 为独立同分布的随机变量序列.

我们记得, 根据第八章第 5 节中的定义, 状态 $x \in E$, 称为

自返的, 如果 $\mathbf{P}_x\{\sigma_x < \infty\} = 1$,
非自返的, 如果 $\mathbf{P}_x\{\sigma_x < \infty\} < 1$.

既然

$$\mathbf{P}_x\{\sigma_x < \infty\} = 1 \quad \Leftrightarrow \quad \mathbf{P}_x\{X_n = x \ \text{ i.o.}\} = 1,$$

$$\mathbf{P}_x\{\sigma_x < \infty\} = 1 \quad \Leftrightarrow \quad \mathbf{P}_x\{X_n = x \ \text{ i.o.}\} = 0$$

(第八章第 5 节定理 1), 所以状态 $x \in E$ 是

自返的, 如果 $\mathbf{P}_x\{X_n = x \ \text{ i.o.}\} = 1$,
非自返的, 如果 $\mathbf{P}_x\{X_n = x \ \text{ i.o.}\} = 0$.

(性质 "$\mathbf{P}_x\{X_n = x \text{ i.o.}\} = 1$" 与 "$\mathbf{P}_x\{X_n = x \text{ i.o.}\} = 0$" 对于诠释 "自返性" 与 "非自返性", 比 "$\mathbf{P}_x\{\sigma_x < \infty\} = 1$" 与 "$\mathbf{P}_x\{\sigma_x < \infty\} < 1$" 更为贴切, 若能将它们分别理解为: "在每次访问状态 x 之后还能再回来" 与 "在某次访问状态 x 之后不再回来" 的话.)

事实上, 状态 x 的自返性等价于下列性质中的任何一个:

$$\mathbf{P}_x\{X_n = x \text{ i.o.}\} = 1 \quad \text{或} \quad \mathbf{P}_x\{N_x = \infty\} = 1 \quad \text{或} \quad \mathbf{E}_x N_x = \infty,$$

而状态 x 的非自返性则等价于下列性质中的任何一个:

$$\mathbf{P}_x\{X_n = x \text{ i.o.}\} = 0 \quad \text{或} \quad \mathbf{P}_x\{N_x < \infty\} = 1 \quad \text{或} \quad \mathbf{E}_x N_x < \infty.$$

- 回到位势理论, 尤其是回到 "再生循环" 概念, 我们就可以对具有可数状态集合的马尔可夫链的不变测度 (或称平稳测度) 与分布 (即概率测度) 的结构问题给出一个详尽的回答了.

我们记得, 对每个转移概率矩阵 $P = \|p_{xy}\|$, $x, y \in E$ 和函数 $f \in \mathscr{E}_+$, 都有一个线性算子 $\mathbb{P}f$ 与它们相联系, 它以如下的方式作用在非负函数 $f = f(x)$, $x \in E$ 上:

$$(\mathbb{P}f)(x) = \sum_{y \in E} p_{xy} f(y), \quad x \in E$$

(即按法则 (矩阵 P)⊗(列向量 f)=(列向量 $\mathbb{P}f$)).

设 $q = q(A)$, $A \subseteq E$ 为可数集合 E 的子集 A 上的非退化的测度 (即不恒等于 0 或无穷). 这样的测度显然完全被它在各个单点集 $\{x\}$, $x \in E$ 上的值 $q(\{x_i\})$ 所确定. (为方便起见, 将 $q(\{x_i\})$ 简写为 $q(x_i)$.)

以 $\mathscr{M}_+$ 表示所有这样的测度 q 的集合, 以 $q\mathbb{P}$ 表示按如下方式将 $\mathscr{M}_+$ 中的测度变为 $\mathscr{M}_+$ 中的测度的线性算子:

$$q\mathbb{P}(y) = \sum_{x \in E} q(y) p_{xy}$$

(即按法则 (行向量 q)⊗(矩阵 P)=(行向量 $q\mathbb{P}$).)

测度 $q \in \mathscr{M}_+$ 称为关于具有算子 $\mathbb{P}$ 的马尔可夫链是不变的 (平稳的), 如果 $q\mathbb{P} = q$. 如果测度 $q \in \mathscr{M}_+$ 使得 $q\mathbb{P} \leqslant q$, 则它称为峰态的, 或 $\mathbb{P}$-峰态的.

在位势理论中, 对于 $f \in \mathscr{E}_+$ 和 $q \in \mathscr{M}_+$ 还引入了如下形式的双线性算子:

$$\langle q, f\rangle = \sum_x q(x) f(x).$$

不难验证, 该双线性算子具有如下所示的对偶性:

$$\langle q, \mathbb{P}f\rangle = \langle q\mathbb{P}, f\rangle,$$

即允许将 "作用在函数上的算子" 变换为 "作用在测度上的算子".

第八章第 6 节定理 2 证明了, 如果 (具有可数状态集合) 的马尔可夫链仅存在一个不可约正定自返类, 那么不变分布存在、唯一, 且有

$$q(x) = [\mathbf{E}_x \sigma_x]^{-1}, \quad x \in E,$$

其中 $\sigma_x = \inf\{n \geqslant 1 : \ X_n = x\}$ 为首次返回状态 x 的时刻. (注意, 此时我们有 $1 \leqslant \mathbf{E}_x \sigma_x < \infty$, $x \in E$.)

正如下面所要证明的, 利用第一个再生循环的特征, 对于任意一个不可约的自返马尔可夫链, 无需假设其状态的正定性, 均可得到其不变集合的存在性与结构.

亦即有如下的断言.

任何具有可数相空间 E 的不可约自返的马尔可夫链 $X = (X_n)_{n\geqslant 0}$ 都有不变测度 $q = q(A)$, $A \subseteq E$, 其中 $q(E) > 0$ 且 $q(E) \neq \infty$("非平凡性"), 且对任何状态 $x \in E$, 都有 $0 < q(x) < \infty$. 本测度除了所乘的常数因子之外是唯一确定的.

为证这一结论, 只需证明测度

$$q^\circ(x) = \mathbf{E}_{x^\circ} \sum_{k=0}^{\sigma_{x^\circ}-1} I_{\{x\}}(X_k),$$

除专门条件 $q^\circ(x^\circ) = 1$ 之外, 是唯一的非平凡不变测度, 其中 $\sigma_{x^\circ} = \inf\{n \geqslant 1 : X_n = x^\circ\}$.

为证该测度 q° 事实上是不变的 (这也就证明了不变测度的存在性), 只需对任何函数 $f \in \mathscr{E}_+$, 证明

$$\langle q^\circ \mathbb{P},\, f\rangle = \langle q^\circ,\, f\rangle.$$

(利用第八章第 2 节第 13 题中推广的严马尔可夫性) 通过下面的一系列运算, 即可证得这一等式的正确性:

$$\begin{aligned}
\langle q^\circ \mathbb{P},\, f\rangle &= \langle q^\circ,\, \mathbb{P}f\rangle = \mathbf{E}_{x^\circ}\left[\sum_{k=0}^{\sigma_{x^\circ}-1} (\mathbb{P}f)(X_k)\right] = \mathbf{E}_{x^\circ}\left[\sum_{k=0}^{\sigma_{x^\circ}-1} \mathbf{E}_{X_k} f(X_1)\right] \\
&= \sum_{k\geqslant 0} \mathbf{E}_{x^\circ}[I_{\{k<\sigma_{x^\circ}\}} \mathbf{E}_{X_k} f(X_1)] = \sum_{k\geqslant 0} \mathbf{E}_{x^\circ}\{I_{\{k<\sigma_{x^\circ}\}} \mathbf{E}_{x^\circ}[f\circ\theta_k|\mathscr{F}_k]\} \\
&= \sum_{k\geqslant 0} \mathbf{E}_{x^\circ}\{\mathbf{E}_{x^\circ}[I_{\{k<\sigma_{x^\circ}\}} f\circ\theta_k|\mathscr{F}_k]\} = \sum_{k\geqslant 0} \mathbf{E}_{x^\circ}[I_{\{k<\sigma_{x^\circ}\}} f\circ\theta_k] \\
&= \mathbf{E}_{x^\circ}\left[\sum_{k\geqslant 0} I_{\{k<\sigma_{x^\circ}\}} f(X_{k+1})\right] = \mathbf{E}_{x^\circ}\left[\sum_{l=1}^{\sigma_{x^\circ}} f(X_l)\right] = \mathbf{E}_{x^\circ}\left[\sum_{k=0}^{\sigma_{x^\circ}-1} f(X_k)\right] \\
&= \langle q^\circ,\, f\rangle.
\end{aligned}$$

所构造出的以 $q^\circ(x^\circ)=1$ 为标志条件的测度 q° 对一切 $x\in E$, 都满足性质 $0<q(x)<\infty$. 关于这一点, 可以由下面的峰态测度的性质直接推出.

设马尔可夫链为不可约的, 而测度 $q\in\mathscr{M}_+$ 是峰态的 ($q\mathbb{P}\leqslant q$). 则若对某个状态 $x^\circ\in E$, 有值 $q(x^\circ)=0$ (值 $q(x^\circ)<\infty$), 那么对一切 $x\in E$, 就都有值 $q(x)=0$ (值 $q(x)<\infty$).

事实上, 对一切 $x\neq x^\circ$, 都存在 $n\geqslant 1$, 使得 $p_{x,x^\circ}^{(n)}>0$, 然而

$$q(x^\circ) \geqslant \sum_{y\in E} q(y) p_{y,x^\circ}^{(n)} \geqslant q(x) p_{x,x^\circ}^{(n)},$$

由此即得所证.

现证测度 q° 是唯一的非平凡的不变测度. 为此, 假设 q 是峰态的 (特别地, 是不变的) 测度, 对一切 $x\in E$, 都有 $0<q(x)<\infty$. 令

$$f(x) = \frac{q(x)}{q^\circ(x)},$$

并定义 (与 $p_{x,y}$ 对偶的) 函数 $\widehat{p}_{x,y} = \frac{q^\circ(y)}{q^\circ(x)} p_{y,x}$. 由于

$$\sum_{y\in E} \widehat{p}_{x,y} = \frac{1}{q^\circ(x)} \sum_{y\in E} q^\circ(y) p_{y,x} = \frac{q^\circ(x)}{q^\circ(x)} = 1,$$

故可见 $\widehat{P}=\|\widehat{p}_{x,y}\|$ 就是转移概率矩阵. 并且, 对于矩阵 $\widehat{P}=\|\widehat{p}_{x,y}\|$ 和函数

$f(x)=\frac{q(x)}{q^\circ(x)}$, 我们有

$$\begin{aligned}
\widehat{\mathbb{P}} f(x) &= \sum_{y\in E} \widehat{p}_{x,y} f(y) = \sum_{y\in E} \widehat{p}_{x,y} \frac{q(y)}{q^\circ(y)} = \sum_{y\in E} \frac{q^\circ(y)}{q^\circ(x)} p_{y,x} \frac{q(y)}{q^\circ(y)} \\
&= \frac{1}{q^\circ(x)} \sum_{y\in E} p_{y,x} q(y) = \frac{q(x)}{q^\circ(x)} = f(x).
\end{aligned}$$

事实上, 函数 $f=f(x)$ 是 $\widehat{\mathbb{P}}$-调和的. 又由于由函数 $\widehat{p}_{x,y}=\frac{q^\circ(y)}{q^\circ(x)}p_{y,x}$ 的定义可以推出: 对任何 $n\geqslant 1$, 都有

$$\widehat{p}_{x,y}^{(n)} = \frac{q^\circ(y)}{q^\circ(x)} p_{y,x}^{(n)},$$

所以

$$\widehat{G}(x,y) = \frac{q^\circ(y)}{q^\circ(x)} G(y,x).$$

由上述两个关系式可以推知: 若具有算子 $\mathbb{P}$ 的马氏链 $X=(X_n)_{n\geqslant 0}$ 是不可约的和自返的, 那么具有算子 $\widehat{\mathbb{P}}$ 的对偶马氏链 $X=(X_n)_{n\geqslant 0}$ 就也是不可约的和自返的. 但若函数 $f=f(x)$ 是 (非负) 峰态 (特别地, 是调和的) 函数, 则对于任何初始分布 π, 序列 $(f(X_n))_{n\geqslant 0}$ 都关于测度 $\widehat{\mathbf{P}}_\pi$ 为非负定上鞅. 这一性质曾经在上面指出过. 在那里我们还提到过, 这样的序列 $\widehat{\mathbf{P}}_\pi$-a.s.地存在极限 $\lim\limits_n X_n$ ($\equiv X_\infty$), 因此, 也就存在极限 $\lim\limits_n f(X_n)$. 若马氏链是不可约与自返的, 则对任二状态 x 与 $y\neq x$, X_n 都无限多次地取值 x, 也无限多次地取值 y. 这就意味着 $f(x)=f(y)$, 从而我们证得了 $f(x)\equiv \text{const}$.

总而言之, 我们已经证得了, 任何别的满足条件 $0<q(x)<\infty$, $x\in E$ 的不变分布 q, 同 q° 的区别仅在于所乘的正的常数.

由所证的结果不难推出前面所提到过的关于不可约正定自返马尔可夫链的如下性质: 唯一的不变概率分布 $q^\circ=(q^\circ(x),\ x\in E)$ 具有结构: $q^\circ(x)=[\mathbf{E}_x\sigma_x]^{-1}, x\in E$.

- 再谈一谈关于具有可数状态集合的马尔可夫链的一些遍历性定理, 亦即当 $n\to\infty$ 时, 形如 $\frac{1}{n}\sum\limits_{k=0}^{n-1} f(X_k)$ 的量, 或者, 更为广泛的, 对于某些类型的函数 f 与 g, 形如

$$\sum_{k=0}^{n-1} f(X_k) \Big/ \sum_{k=0}^{n-1} g(X_k)$$

的量的几乎必然收敛性的定理. 与上面一样, 此处最好是考察再生循环.

设 $X=(X_n)_{n\geqslant 0}$ 是不可约自返的马尔可夫链, 具有可数状态集合 E 和这样的不变测度 $q^\circ(x)$: 对一切 $x\in E$, 都有 $0<q^\circ(x)<\infty$, 且对某个固定的 x°, 有 $q^\circ(x^\circ)=1$.

设函数 $f = f(x)$ 与 $g = g(x)$ 均属于类 $L^1(q^\circ)$, 即有 $\sum\limits_{x\in E} |f(x)|q^\circ(x) < \infty$ 与 $\sum\limits_{x\in E} |g(x)|q^\circ(x) < \infty$. 令

$$Y_0 = \sum_{k=0}^{\sigma_{x^\circ}^1 - 1} f(X_k), \quad 及 \quad Y_m = \sum_{k=\sigma_{x^\circ}^m}^{\sigma_{x^\circ}^{m+1}-1} f(X_k) \quad (= Y_0 \circ \theta_{\sigma_{x^\circ}^m}).$$

根据不变测度 q° 的定义, 有

$$\mathbf{E}_x Y_0 = \langle q^\circ, f \rangle,$$

且 (根据马尔可夫性) 对任何初始分布 π, 都有

$$\mathbf{E}_\pi Y_m = \mathbf{E}_\pi[\mathbf{E}_{X_{\sigma_{x^\circ}^m}}(Y_0)] = \mathbf{E}_{x^\circ} Y_0 = \langle q^\circ, f \rangle.$$

事实上, 关于测度 $\mathbf{P}_\pi$, 随机变量 $Y_1, Y_2, \cdots$ 是独立同分布的, 有 $\mathbf{E}_\pi Y_m = \langle q^\circ, f \rangle$ $(< \infty)$. 所以, 由强大数律可知, (对于任何 π, 都 $\mathbf{P}_\pi$-a.s.地) 有

$$\frac{1}{n}\sum_{k=0}^{\sigma_{x^\circ}^n} f(X_k) = \frac{Y_0}{n} + \frac{1}{n}(Y_1 + \cdots + Y_{n-1}) + \frac{f(x^\circ)}{n} \to \langle q^\circ, f \rangle, \quad n \to \infty,$$

又在 $\langle q^\circ, g \rangle \neq 0$ 的假设之下, (对于任何 π, 都 $\mathbf{P}_\pi$-a.s.地) 有

$$\frac{\sum\limits_{k=0}^{\sigma_{x^\circ}^n} f(X_k)}{\sum\limits_{k=0}^{\sigma_{x^\circ}^n} g(X_k)} \to \frac{\langle q^\circ, f \rangle}{\langle q^\circ, g \rangle} \quad (\mathbf{P}_\pi\text{-a.s.}), \quad n \to \infty.$$

若记 $\nu_{x^\circ}^n = \sum\limits_{k=0}^{n} I(X_k = x^\circ)$, 则由马氏链的自返性知 $\nu_{x^\circ}^n \to \infty$ $(\mathbf{P}_\pi\text{-a.s.})$, $n \to \infty$. 由于 $\sigma_{x^\circ}^{\nu_{x^\circ}^n} \leqslant n < \sigma_{x^\circ}^{\nu_{x^\circ}^n + 1}$, 故由上述收敛性即可得出关于比值的遍历性定理: 当 $n \to \infty$ 时, 有

$$\frac{\sum\limits_{k=0}^{n} f(X_k)}{\sum\limits_{k=0}^{n} g(X_k)} \to \frac{\langle q^\circ, f \rangle}{\langle q^\circ, g \rangle} \quad (\mathbf{P}_\pi\text{-a.s.}).$$

假若所考察的马氏链是不可约正定自返的, 那么由于此时可将测度 q° 换为概率分布 $\pi^\circ = (\pi^\circ(x),\ x \in E)$, 其中 $\pi^\circ(x) = 1/(\mathbf{E}_x \sigma_x)$, 所以我们可以特别地得到关于这样的马氏链的如下形式的遍历性定理:

$$\frac{1}{n}\sum_{k=0}^{n} f(X_k) \to \langle \pi^\circ, f \rangle \quad (\mathbf{P}_\pi\text{-a.s.}), \quad n \to \infty,$$

其中 π 为任意的初始分布 (特别地, 可将其取为 π°).

参考文献

[1] Адамс, Гилемин (Adams M., Guillemin V.). Measure Theory and Probability. — Monterey, CA: Wadsworth & Brooks/Cole, 1986.

[2] Альдус (Aldous D. J.). Exchangeability and Related Topics. — Berlin: Springer-Verlag, 1985. — (Lecture Notes in Mathematics. — V. 1117.)

[3] Арнольд В. И. Цепные дроби. — М.: МЦНМО, 2001.

[4] Арнольд В. И. Что такое математика? — М.: МЦНМО, 2004.

[5] Баранов В. И., Стечкин Б. С. Экстремальные комбинаторные задачи и их приложения. — М.: Физматлит, 2004.

[6] Бенткус (Bentkus V.). On Hoeffding's inequalities. — Annals of Probability. — 2004. — V. 32, № 2. — P. 1650—1673.

[7] Бергер (Berger M. A.). An Introduction to Probability and Stochastic Processes. — New York: Springer-Verlag, 1993.

[8] Бертран (Bertrand J.). Calcul des Probabilités. — Paris, 1889.

[9] Биллингсли П. Сходимость вероятностных мер. — М.: Наука, 1977.

[10] Биллингсли (Billingsley P.). Probability and Measure. — 3rd ed. — New York: Wiley, 1995.

[11] Болдин М. В., Симонова Г. И., Тюрин Ю. Н. Знаковый статистический анализ линейных моделей. — М.: Наука, 1997.

[12] Большев Л. Н., Смирнов Н. В. Таблицы математической статистики. — М.: Наука, 1983.

[13] Бремо (Brémaud P.). An Introduction to Probabilistic Modeling. — New York: Springer-Verlag, 1988.

[14] Бремо (Brémaud P.). Markov Chains. Gibbs Fields, Monte Carlo Simulation, and

Queues. — New York: Springer-Verlag, 1999.

[15] Брэдли (Bradley R. C.). The central limit question under ρ-mixing. — Rocky Mountain Journal of Mathematics. — 1987. — V. 17, № 1. — P. 95 — 114; corrections: *ibid.* — № 4. — P. 891.

[16] Булинский А. В. Предельные теоремы в условиях слабой зависимости. — М.: Изд-во Моск. ун-та, 1989.

[17] Булинский А. В., Ширяев А. Н. Теория случайных процессов. — М.: Физматлит, 2003.

[18] Буркхольдер (Burkholder D. L.). Sufficiency in the undominated case. — Ann. Math. Statist. — 1961. — V. 32, № 4. — P. 1191—1200.

[19] Бхаттачарья, Веймир (Bhattacharya R. N., Waymire E. C.). Stochastic Processes with Applications. — New York: Wiley, 1990.

[20] Ватанабэ С., Икэда Н. Стохастические дифференциальные уравнения и диффузионные процессы. — М.: Наука, 1986.

[21] Вероятность и математическая статистика. Энциклопедия / Гл. ред. Ю. Прохоров. — Научное изд-во《Большая российская энциклопедия》, 1999.

[22] Вильямс (Williams D.). Probability with Martingales. — Cambridge: Cambridge Univ. Press, 1991.

[23] Вильямс (Williams D.). Weighing the Odds. A Course in Probability and Statistics. — Cambridge: Cambridge Univ. Press, 2001.

[24] Гливенко В. И. Теория вероятностей. М.: — Учпедгиз, 1937.

[25] Гмурман В. Е. Руководство к решению задач по теории вероятностей и математической статистике. — М.: Высшая школа, 1979.

[26] Гнеденко Б. В. Курс теории вероятностей. — М.: УРСС, 2001.

[27] Гончаров В. Л. Теория вероятностей. — М. — Л.: Гос. изд-во оборонной промышл., 1939.

[28] Градштейн И. С., Рыжик И. М. Таблицы интегралов, сумм, рядов и произведений. — М.: Физматгиз, 1962.

[29] Гримметт, Стирзакер (Crimmett G. R., Stirzaker D. R.). One Thousand Exercises in Probability. Companion to Probability and Random Processes. — Oxford: Oxford Univ. Press, 2001.

[30] Гут (Gut A.). Stopped Random Walks. — New York: Springer-Verlag, 1988.

[31] Дарретт (Durrett R.). Probability: Theory and Examples. — Pacific Grove, CA: Wadsworth & Brooks/Cole, 1991.

[32] Де Гроот, Шервиш (De Groot M. H., Schervish M. J.) Probability and statistics. — Addison-Wesley, 2002.

[33] де Финетти (de Finetti B.). Theory of Probability: A Critical Introductory Treatment. — V. 1, 2. London, etc.: Wiley, 1974, 1975.

[34] Деврой (Devroye L.). A Course in Density Estimation. — Boston: Birkhäuser, 1987.

[35] Дороговце в А. Я., Сильвестров Д. С., Скороход А. В., Ядренко М. И. Теория вероятностей (сборник задач). — Киев: Вища школа, 1980.

[36] Дэйвид Г. Порядковые статистики. — М.: Наука, 1979.

[37] Дьяченко М. И., Ульянов П. Л. Мера и интеграл. — М.: Факториал, 2002.

[38] Емельянов Г. В., Скитович В. П. Задачник по теории вероятностей и математической статистике. — Л.: Изд-во Ленингр. ун-та, 1967.

[39] Жакод, Проттер (Jacod J., Protter Ph.). Probability Essentials. — Berlin: Springer-Verlag, 2003.

[40] Ибрагимов И. А., Розанов Ю. А. Гауссовские случайные процессы. — М.: Наука, 1970.

[41] Ивченко Г. И., Медведев Ю. И. Математическая статистика. — М.: Высшая школа, 1984.

[42] Имхоф (Imhof J.-P.). Introduction au Calcul des Probabilités. — Paris: Gauthier-Villars, 1969.

[43] Кавата (Kawata T.). Fourier Analysis in Probability Theory. — New York — London: Academic Press, 1972.

[44] Какуллос (Cacoullos T.). Exercises in Probability. — New York: Springer-Verlag, 1989.

[45] Камерон (Cameron P. J.). Combinatorics. Topics, Techniques, Algorithms. Cambridge: Cambridge Univ. Press., 1994.

[46] Капиньский, Заставняк (Capiński M., Zastawniak T.). Probability through Problems. — New York: Springer-Verlag, 2001.

[47] Карлин, Тейлор (Karlin S., Taylor H. M.). A First Course in Stochastic Processes. — New York — London: Academic Press, 1975.

[48] Карлин, Тейлор (Karlin S., Taylor H. M.). A Second Course in Stochastic Processes. — New York — London: Academic Press, 1981.

[49] Карр (Karr A. F.). Probability. — Berlin etc.: Springer-Verlag, 1993.

[50] Кемени, Снелл (Kemeny J. G., Snell J. L.). Finite Markov Chains. — Princeton, NJ, etc.: Van Nostrand, 1960.

[51] Кемени Д ж., Снелл Д ж., Томпсон Д ж. Введение в конечную математику. — М.: ИЛ, 1963.

[52] Кендалл М., Моран П. Геометрические вероятности. — М.: Наука, 1972.

[53] Кеннан (Kannan D.). An Introduction to Stochastic Processes. — New York; Oxford: North-Holland, 1979.

[54] Керубини, Лучано, Веккьято (Cherubini D., Luciano E., Vecchiato W.). Copula Methods in Finance. — New York: Wiley, 2004.

[55] Кингман (Kingman J. F. C.). Random variables with unsymmetrical linear regressions // Mathematical Proceedings of the Cambridge Philosophical Society. — 1985. — V. 98, № 2. — P. 355—365.

[56] Коваленко И. Н., Сарманов О. В. Краткий курс теории случайных процессов.

— Киев: Вища школа, 1978.
[57] Козлов М. В. Элементы теории вероятностей в примерах и задачах. — М.: Изд-во Моск. ун-та, 1990.
[58] Козлов М. В., Прохоров А. В. Введение в математическую статистику. — М.: Изд-во Моск. ун-та, 1987.
[59] Колчин В. Ф. Случайные отображения. — М.: Наука, 1984.
[60] Колмогоров А. Н. Основные понятия теории вероятностей. — 3-е изд. М.: ФАЗИС, 1998.
[61] Константин (Constantine G. M.). Combinatorial Theory and Statistical Design. — New York: Wiley, 1987.
[62] Копп (Kopp P. E.). Martingales and Stochastic Integrals. — Cambridge: Cambridge Univ. Press, 1984.
[63] Корнфельд И. П., Синай Я. Г., Фомин С. В. Эргодическая теория. — М.: Наука, 1980.
[64] Коц, Надарая (Kotz S., Nadarajah S.). Extreme Value Distributions. Theory and Applications. — London: Imperial College Press, 2000.
[65] Кочен, Стоун (Kochen S., Stone Ch.). A note on the Borel — Cantelli lemma // Illinois Journal of Mathematics. — 1964. — V. 8. — P. 248— 251.
[66] Криф, Леви (Krief A., Levy S.). Calcul des Probabilités. Exercises. — Paris: Hermann, 1972.
[67] Ландо С. К. Лекции о производящих функциях. — М.: МЦНМО, 2002.
[68] Леман Э. Проверка статистических гипотез. М.: Наука, 1964.
[69] Летак (Letac G.). Problèmes de Probabilité. Paris: Presses Universitaires de France, 1970.
[70] Линдгрен, Мак-Элрат (Lindgren B. W., McElrath G. W.). Introduction to Probability and Statistics. — London — New York: Macmillan, 1966.
[71] Липцер Р. Ш., Ширяев А. Н. Статистика случайных процессов. М.: Наука, 1974.
[72] Липцер Р. Ш., Ширяев А. Н. Теория мартингалов. М.: Наука, 1986.
[73] Ловас, Пеликан, Вестергомби (Lovász L., Pelicán J., Vesztergombi K.). Discrete Mathematics. Elementary and Beyond. — Springer, 2003.
[74] Лон (Long R. L.). Martingale Spaces and Inequalities. — Beijing; Braunschweig: Peking Univ. Press; Vieweg, 1993.
[75] Лоулер (Lawler G. F.). Intersections of Random Walks. — Boston: Birkhäuser, 1991.
[76] Лукач Е. Характеристические функции. — М.: Наука, 1979.
[77] Лукач (Lukacs E.). Developments in Characteristic Function Theory. — New York: Macmillan, 1983.
[78] Мазлиак, Приуре, Бальди (Mazliak L., Priouret P., Baldi P.). Martingales et chaînes de Markov. — Paris: Hermann, 1998.

[79] Математическая энциклопедия: В 5 т. / Гл. ред. И. М. Виноградов. М.: Советская энциклопедия, 1977—1985.
[80] Местер (Meester R.). A Natural Introduction to Probability Theory. — Basel: Birkhäuser, 2003.
[81] Микош (Mikosch T.). Non-Life Insurance Mathematics. — Berlin etc.: Springer-Verlag, 2004.
[82] Моран (Moran P. A. P.). An Introduction to Probability Theory. — Oxford: Clarendon Press, 1968.
[83] Мостеллер Ф., Рурке Р., Томас Д ж. Вероятность. — М.: Мир, 1969.
[84] Невзоров В. Б. Рекорды: Математическая теория. — М.: ФАЗИС, 2000.
[85] Норрис (Norris J. R.). Markov Chains. — Cambridge: Cambridge Univ. Press, 1999.
[86] Ортега, Вшебор (Ortega J., Wschebor M.). On the sequence of partial maxima of some random sequences. — Stochastic Processes and their Applications. — 1984. — V. 16, № 1. — P. 85—98.
[87] Петров (Petrov V. V.). A generalization of the Borel — Cantelli lemma // Statistics and Probability Letters. — 2004. — V. 67, № 3. — P. 233—239.
[88] Петров В. В. Предельные теоремы для сумм независимых случайных величин. — М.: Наука, 1987.
[89] Летров, Мордецкий (Petrov V., Mordecki E.). Theoría de Probabilidades. (Con más de 200 ejercicios.) — Moscow: Editorial URSS, 2002.
[90] Пешкир Г., Ширяев А. Н. Неравенства Хинчина и мартингальное расширение сферы их действия. — Успехи математических наук. — 1995. — Т. 50, вып. 5. — С. 3—62.
[91] Порт (Port S. C.) Theoretical Probability for Applications. — New York: Wiley, 1994.
[92] Прохоров А. В., Ушаков В. Г., Ушаков Н. Г. Задачи по теории вероятностей: Основные понятия. Предельные теоремы. Случайные процессы. — М.: Наука, 1986.
[93] Пуанкаре (Poincaré H.). Calcul des probabilités. — 2éd. — Paris, 1912. Рус. пер.: Пуанкаре А. Теория вероятностей. — Ижевск: Ред. журн. 《Регулярная и хаотическая динамика》, 1999.
[94] Ревес (Révész P.). The Law of Large Numbers. — New York — London: Academic Press, 1968.
[95] Ревес (Révész P.). Random Walk in Random and Non-random Environments. — Teaneck, N. J.: World Scientific, 1990.
[96] Ревес, Тот (Révész P., Tóth B., Eds.). Random Walks. — Budapest: János Bolyai Mathematical Society, 1999. — (Bolyai Society Mathematical Studies. — V. 9.)
[97] Ревуз, Йор (Revuz D, Yor M.). Continuous Martingales and Brownian Motion. — 3rd ed. — Berlin: Springer-Verlag, 1999. Рус. пер.: Непрерывные мартингалы и броуновское движение. М.: МЦНМО (в печати).

[98] Розенталь (Rosenthal J. S.). A First Look at Rigorous Probability Theory. — River Edge, N. J.: World Scientific, 2000.

[99] Романо, Зигель (Romano J. P., Siegel A. F.). Counterexamples in Probability and Statistics. — Monterey, CA: Wadsworth & Brooks/Cole, 1986.

[100] Росс (Ross S. M.). Introduction to Probability Models. — 8th ed. Boston: Academic Press, 2003.

[101] Росс (Ross S. M.). Stochastic Processes. — 2nd ed. New York: Wiley, 1996.

[102] Ружа, Секей (Ruzsa I. Z., Székely G. J.). Algebraic Probability Theory. Chichester: Wiley, 1988.

[103] Сачков В. Н. Комбинаторные методы дискретной математики. — М.: Наука, 1977.

[104] Сачков В. Н. Вероятностные методы в комбинаторном анализе. — М.: Наука, 1978.

[105] Сборник задач по теории вероятностей, математической статистике и теории случайных функций / Под ред. А. А. Свешникова. — М.: Наука, 1965.

[106] Севастьянов Б. А. Курс теории вероятностей и математической статистики. — М. — Ижевск: Институт компьютерных исследований, 2004.

[107] Севастьянов Б. А., Чистяков В. П., Зубков А. М. Сборник задач по теории вероятностей. — М.: Наука, 1980.

[108] Секей Г. Парадоксы в теории вероятностей и математической статистике. — 2-е изд. М. — Ижевск: Институт компьютерных исследований, 2003.

[109] Сешадри (Seshadri V.). The Inverse Gaussian Distribution. — New York: Clarendon Press, 1993.

[110] Скляр (Sklar A.). Fonctions de répartition à n dimensions et leurs marges // Publications de l'Institut de Statistique de l'Université de Paris. — 1959. — V. 8. — P. 229—231.

[111] Скляр (Sklar A.). Random variables, distribution functions, and copulas — a personal look backward and forward // IMS Lecture Notes — Monograph Series. — 1996. — V. 28. P. 1—14.

[112] Скороход А. В. Вероятность вокруг нас. — Киев: Наукова думка, 1980.

[113] Спицер Ф. Принципы случайного блуждания. — М.: Мир, 1969.

[114] Стоянов, Мирачийски, Игнатов, Танушев (Stoyanov J., Mirazchiiski I., lgnatov Z., Tanushev M.). Exercise Manual in Probability Theory. — Dordrecht: Kluwer, 1989.

[115] Стоянов Й. Контрпримеры в теории вероятностей. — М.: Факториал, 1999.

[116] Тутубалин В. Н. Теория вероятностей и случайных процессов. — М.: Изд-во Моск. ун-та, 1992.

[117] Уайз, Холл (Wise G. L., Hall E. B.). Counterexamples in Probability and Real Analysis. — New York: Clarendon Press, Oxford Univ. Press, 1993.

[118] Уолл (Wall C. R.). Terminating decimals in the Cantor ternary set // Fibonacci Quarterly. — 1990. — V. 28, №2. — P. 98—101.

[119] Феллер В. Введение в теорию вероятностей и ее приложения: В 2-х т. — М.: Мир, 1984.

[120] Фристедт, Грей (Fristedt B., Gray L.). A Modern Approach to Probability Theory. — Boston: Birkhäuser, 1997.

[121] Хоффман-Йёнсен (Hoffmann-Jørgensen J.). Probability with a View toward Statistics. V. I, II. New York: Chapman & Hall, 1994.

[122] Холл М. Комбинаторика. — М.: Мир, 1970.

[123] Хренников (Khrennikov A.). Interpretations of Probability. — Utrecht: VSP, 1999.

[124] Чандрасекар С. Стохастические проблемы в физике и астрономии. — М.: ИЛ, 1947.

[125] Чибисов Д. М., Пагурова В. И. Задачи по математической статистике. — М.: Изд-во Моск. ун-та, 1990.

[126] Чжун Кай-лай. Однородные цепи Маркова. — М.: Мир, 1964.

[127] Шафер, Вовк (Shafer G., Vovk V.). Probability and Finance. — New York: Wiley, 2001.

[128] Шервиш (Schervish M. J.). Theory of Statistics. — New York: Springer-Verlag, 1995.

[129] Ширяев А. Н. Статистический последовательный анализ. — М.: Наука, 1976.

[130] Ширяев А. Н. Основы стохастической финансовой математики: В 2-х т. — М.: ФАЗИС, 1998 (1-е изд.), 2004 (2-е изд.).

[131] Шомон, Йор (Chaumont L., Yor M.). Exercises in Probability. — Cambridge: Cambridge Univ. Press, 2003.

[132] Эверитт (Everitt B. S.). Chance Rules. An Informal Guide to Probability, Risk, and Statistics. New York: Copernicus, Springer-Verlag, 1999.

[133] Эрдёш П., Спенсер Д ж. Вероятностные методы в комбинаторике. — М.: Мир, 1976.

名词索引

[1] 3.3.15 表示第三章第 3 节第 15 题, 注表示在该题的注释中 —— 译者注.